Robust variability of grid cell properties within individual grid modules enhances encoding of local space

William T. Redman[1,2] †, Santiago Acosta-Mendoza[1] †, Xue-Xin Wei[3,4,5,6], and Michael J. Goard[7,8,9]

[1]Interdepartmental Graduate Program in Dynamical Neuroscience, University of California, Santa Barbara
[2]Intelligent Systems Center, Johns Hopkins University Applied Physics Lab
[3]Department of Neuroscience, The University of Texas at Austin
[4]Department of Psychology, The University of Texas at Austin
[5] Center for Perceptual Systems, The University of Texas at Austin
[6]Center for Theoretical and Computational Neuroscience, The University of Texas at Austin
[7]Department of Psychological and Brain Sciences, University of California, Santa Barbara
[8]Department of Molecular, Cellular, and Developmental Biology, University of California, Santa Barbara
[9]Neuroscience Research Institute, University of California, Santa Barbara
†Equal contribution

Abstract

Although grid cells are one of the most well studied functional classes of neurons in the mammalian brain, whether there is a single orientation and spacing value per grid module has not been carefully tested. We analyze a recent large-scale recording of medial entorhinal cortex to characterize the presence and degree of heterogeneity of grid properties within individual modules. We find evidence for small, but robust, variability and hypothesize that this property of the grid code could enhance the encoding of local spatial information. Performing analysis on synthetic populations of grid cells, where we have complete control over the amount heterogeneity in grid properties, we demonstrate that grid property variability of a similar magnitude to the analyzed data leads to significantly decreased decoding error. This holds even when restricted to activity from a single module. Our results highlight how the heterogeneity of the neural response properties may benefit coding and opens new directions for theoretical and experimental analysis of grid cells.

The discovery of grid cells in medial entorhinal cortex (MEC) [1] has led to considerable experimental and computational work aimed at identifying their origin [2–12] and their function [13–21]. The organization of grid cells into discrete modules [22, 23], with grid properties (grid spacing and orientation) clustered within module, but not between modules, has fundamentally shaped this research. For instance, the increasing size and spacing of grid modules along the dorsal-ventral axis of MEC, by discontinuous jumps of a near constant ratio, has been argued to be optimal for encoding local spatial information [24, 25] when grid cell activity across all modules is integrated together [24–28]. This is despite the fact that the individual neurons in hippocampus, a downstream target of the MEC, receive inputs from only a portion of the dorsal-ventral axis [29]. The modularity of the grid system has also been proposed to simplify the wiring necessary for generating continuous attractor dynamics [2, 3, 5, 16], a computational mechanism theorized to underlie grid cells function that enjoys considerable experimental support [23, 30–34].

Much of this prior theoretical work has made the additional assumption that grid cell properties are identical, up to a phase shift, within a single module [24, 25, 27, 28, 35]. However, as experimentalists have been aware, the distributions of measured orientation and spacing show non-zero variability [22, 30, 34]. This could be due to finite recording time, neurophysiological noise, and/or the sensitivity of numerical methods used to fit grid properties. Alternatively, this variability could reflect underlying inhomogeneity, at a fine scale, within modules. Despite the fundamental way in which grid cell function has been informed by their modular organization, to the best of our knowledge, no characterization of the degree and robustness of variability in grid properties within individual modules has been performed.

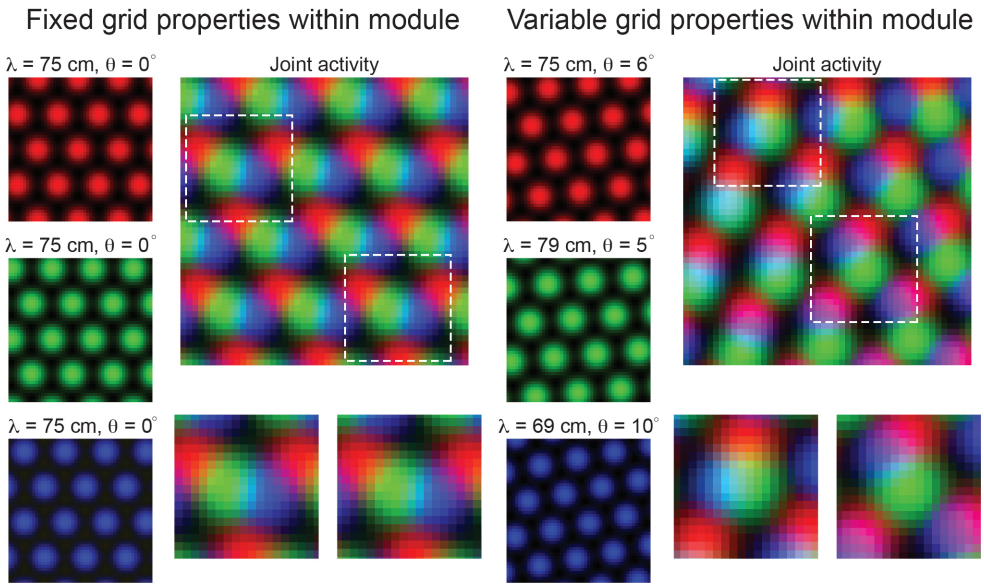

Figure 1: **Variability in grid cell properties within a module leads to enhanced encoding of local space.** When the activity of three idealized grid cells, all with the same grid spacing and orientation, are considered, the periodicity of the responses limits the amount of information conveyed about local space (Left column – "Fixed grid properties within module"). That is, there are multiple locations in physical space with identical population level activity. However, when three grid cells with variable grid spacing and orientation (in the realm of what is measured within individual grid modules – see Results), their joint activity contains considerably more information (Right column – "Variable grid properties within module"). This benefit of spatial inhomogeneity is expected to increase with larger populations of grid cells. Dashed squares in the joint activity map are enlarged below.

If robust variability of grid properties does exist, then it is possible that this heterogeneity could be exploited to encode additional information about local space, reducing the requirement of integration across multiple grid modules (Fig. 1). In particular, when all grid cells in a module have the same grid orientation and spacing, the joint population activity has a translational invariance that persists, even as larger populations of grid cells are considered (Fig. 1, "Fixed grid properties within module"). Thus, it is not possible to resolve spatial position from population activity. In contrast, if grid cells within a module have variability in their grid orientation and spacing, then – over a finite area – the translational invariance of the population activity is broken. In such a case, distinct patterns of population activity emerge in distinct locations of space (Fig. 1, "Variable grid properties within module"). As an example, the blue and green grid cells in Fig. 1B show the most overlap in the upper left half of the arena (denoted by cyan pixels), while the red and blue grid cells show the most overlap in the lower right portion of the arena (denoted by purple pixels). This is despite the fact that the variability in spacing and orientation is only on the order of a few centimeters and degrees, respectively. We expect that these patterns should become more complex and contain more information that could be used to disentangle spatial location as more grid cells are considered.

In this paper, we perform detailed analysis of recent state-of-the-art MEC electrophysiological recordings, which were made publicly available [34]. We characterize the variability of grid orientation and spacing within individual modules and find evidence for small, but robust, variability. This variability is present whether a single value of grid orientation and spacing is assigned to each grid cell, or whether grid distortion [36, 37] is taken into account by considering three grid orientations and spacings independently. Performing similar analysis on recent normative recurrent neural network (RNN) models of grid cells [10], we find the presence of comparable variability. This provides another possible explanation for why these RNN models have been found to capture MEC grid cell response profiles [38]. To assess the functional implications of this heterogeneity, we perform simulation experiments with synthetic, noisy grid cell populations, where we have complete control over the distribution of grid orientation and spacing. We find that the variability

in grid cell orientation and spacing, at a similar degree as present in the data we analyze, leads to lower decoding error of local space when using the activity of a *single* module.

Taken together, our results challenge a frequently made assumption in the theoretical literature and support a growing understanding of the spatial information encoded by grid cell populations [39–42]. Additionally, our results encourage consideration of the broader benefits that multiple modules may provide, beyond the encoding of local space [43–45].

Results

Robust differences in grid cell properties within individual modules

To determine the extent to which variability of grid properties in individual modules exists, and to what extent this variability is a robust property of the grid code, we analyzed previously published MEC recordings [34], which include tens to hundreds of grid cells simultaneously recorded. This allows us to characterize the distribution of grid properties within a single grid module, to an extent not possible with other data sets.

For each grid cell, we compute the grid spacing (λ) and orientation (θ) by measuring properties associated with the six hexagonally distributed peaks of the spatial autocorrelogram (SAC), as traditionally performed [46] (Fig. 2A; see Materials and Methods). For clarity, we begin by focusing on a single module, recorded from an open field environment (recording identifier: Rat R, Day 1, Module 2 – R12). This module was picked for its long recording time (approximately 130 minutes of recorded activity, as compared to the other open field recordings that have 30 minutes or less of recorded activity) and for its large number of simultaneously recorded putative grid cells ($N = 168$). In this module, we find that grid cells with high grid scores (> 0.85; $N = 74$) have a range of grid orientation and spacing (example cells shown in Fig. 2B; distributions across module shown in Fig. 2E, F), with λ ranging from 65 cm to 90 cm and θ ranging from 2° to 9°. Overlaying the SACs of pairs of grid cells with similar grid spacing and different grid orientation (Fig. 2C) or vice versa (Fig. 2D) enables visualization of the extent of this variability.

Because individual grid fields can be cut-off by the boundaries of the environment, it is possible that computing the grid spacing from the SAC (which considers all grid fields) could lead to an under-estimate of λ for some grid cells. To verify that the broad distribution of grid spacing, within the same module, that we see is not due to the specifics of the SAC, we recomputed the grid spacing for all grid cells directly from their rate maps, considering only the three grid fields closest to the center of the environment (Fig. S1A; see Materials and Methods). We find that while the grid spacing estimated from the SAC tends to be larger than the grid spacing estimated from the rate maps, a similarly broad range of λ is again present (Fig. S1B, C).

To assess whether the heterogeneity of grid properties present in a single grid module is a robust feature or attributable to noise (either in the recording or the grid property fitting procedure), we measure the variability in grid orientation and spacing **within** a single grid cell and **between** pairs of grid cells. If the heterogeneity is explainable by noise, then we expect that the within-cell variability will be of the same magnitude as the between-cell variability. In contrast, if the heterogeneity is a robust feature of the grid code, then we expect the within-cell variability will be significantly smaller than the between-cell variability.

To measure the within- and between-cell variability, we split the recording into evenly spaced 30 second bins, randomly assigning each temporal bin to one of two equal length halves and computing the grid properties for each half of the recording (Fig. 3A; see Materials and Methods). We set inclusion criteria to filter out cells that do not have consistent hexagonally structured SACs across splits of the recording (see Materials and Methods). Although these requirements are strict (see Fig. S2 for percent of cells rejected), they set a conservative estimate on the amount of grid property variability, ensuring that we do not artificially inflate the variability due to inclusion of unreliable grid cells. We found that the length of the temporal bin used to split the data does not have a large impact on the percentage of cells accepted by this criteria (Fig. S3; see Materials and Methods).

We find that the distribution, across 100 random shuffles of the data into two halves, of within-cell

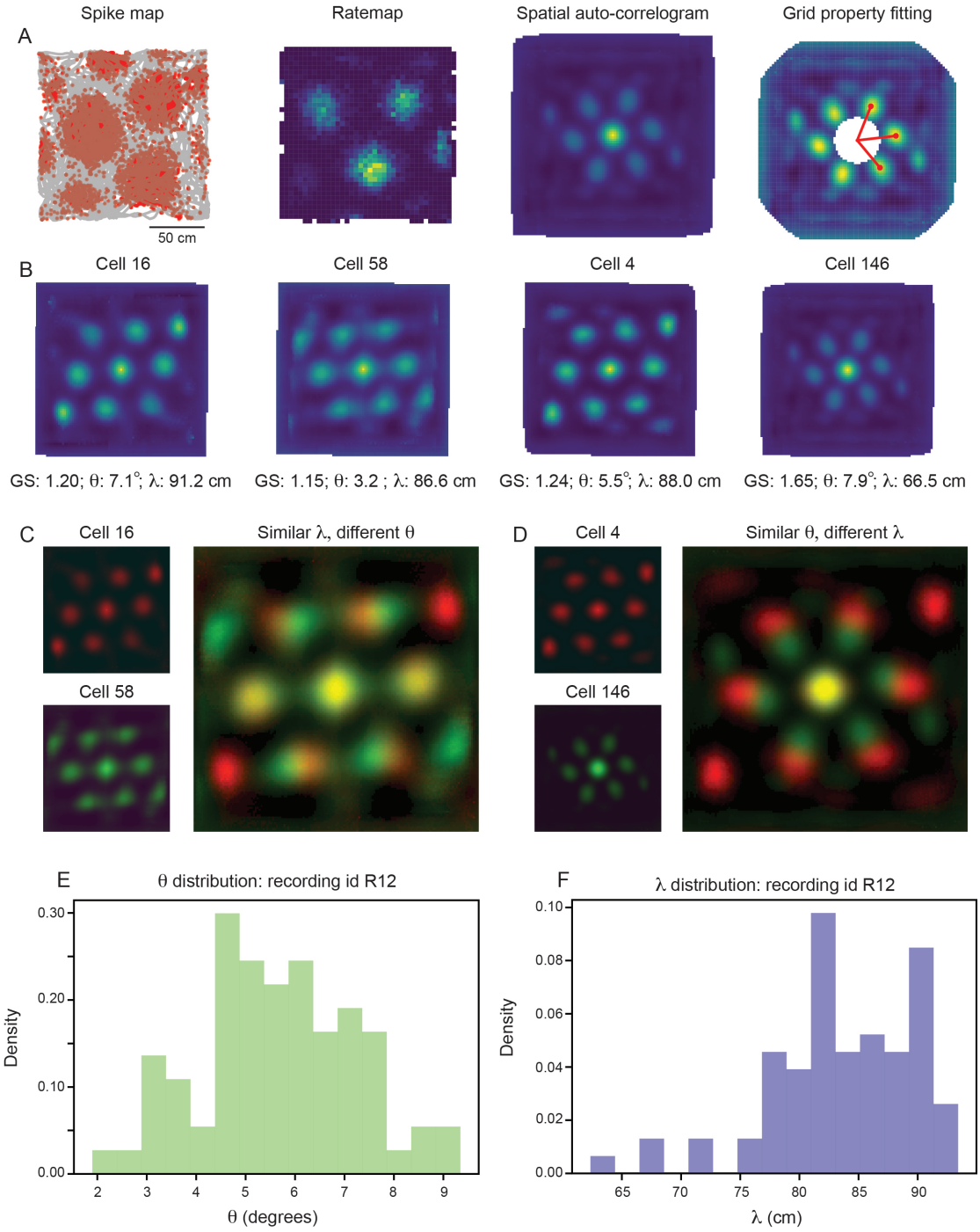

Figure 2: **Grid properties are variable within a single grid module (recording ID R12).** (A) Overview of the standard procedure used to calculate the grid spacing and orientation of a given grid cell. First, spike maps are computed by identifying the location of the animal at the time of each spike. Gray line denotes the trajectory of the rat, red dots denote locations of spikes. A rate map is constructed by binning space and normalizing by the amount of time the rat spent in each spatial bin. A spatial autocorrelogram (SAC) is computed and, after the center peak is masked out (white pixels in the center of the spatial autocorrelogram – leading to change in color scale), the grid properties are fit by measuring the length and angle of the three peaks closest to 0°. (B) Example grid cells from the same module (recording ID R12), with estimated grid score, orientation (θ), and spacing (λ). (C)–(D) SAC overlaid for two pairs of grid cells [from (B)]; one pair with different θ and similar λ (C) and the other with similar θ and different λ (D). (E)–(F) Distribution of θ (E) and λ (F) across all grid cells with grid score > 0.85 (N = 74).

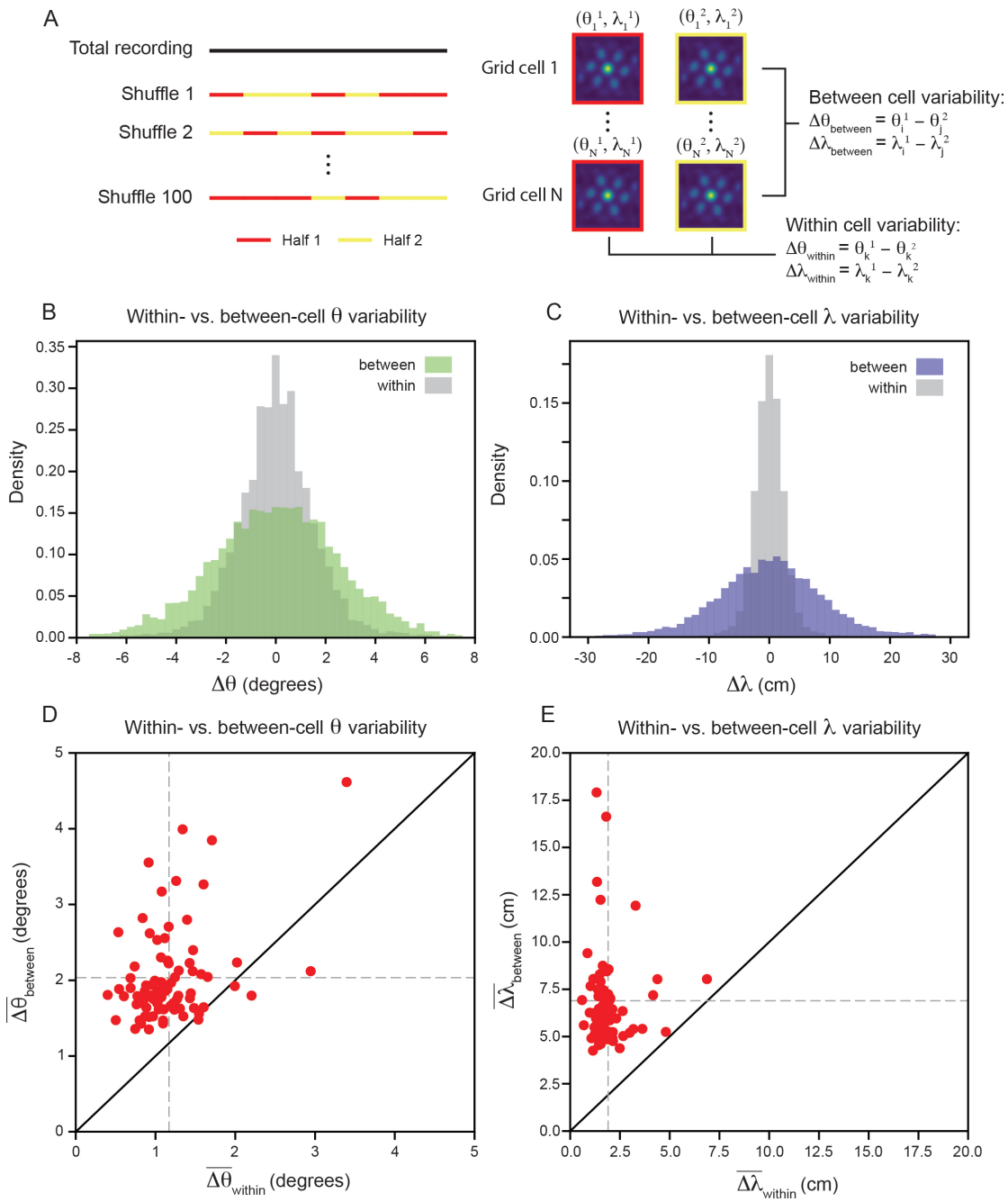

Figure 3: **Variability of grid properties is a robust feature of individual grid module (recording ID R12).** (A) Schematic overview of approach used to compute the between- and within-cell variability of grid orientation and spacing. (B)–(C) Distribution of within- and between-cell variability of θ and λ, respectively. Note that the distribution is across all 100 random shuffles of the data into two halves. (D) Average within-cell variability of grid orientation ($\overline{\Delta\theta}_{\text{within}}$), compared to average between-cell variability of grid spacing ($\overline{\Delta\theta}_{\text{between}}$). (E) Same as (D), but for λ. 1 cell was excluded from (E) for visualization ($\overline{\Delta\lambda}_{\text{between}} > 20$ cm), but was included in non-parametric statistical analysis. For (D)–(E), $N = 82$.

variability of grid orientation and spacing is more concentrated around 0 than the between-cell variability. Comparing the average within- and between-cell variability of grid spacing and orientation reveals that nearly all of the grid cells that passed the criteria for inclusion ($N = 82$) exhibit more between-cell than within-cell variability (Fig. 3D, E): 95.1% grid cells for orientation ($\langle\overline{\Delta\theta}_{\text{within}}\rangle = 1.2°$, $\langle\overline{\Delta\theta}_{\text{between}}\rangle = 2.0°$; $\langle\cdot\rangle$ denotes mean across cells) and 100% of grid cells for spacing ($\langle\overline{\Delta\lambda}_{\text{within}}\rangle = 1.9$ cm, $\langle\overline{\Delta\lambda}_{\text{between}}\rangle = 6.9$ cm). A Wilcoxon-Signed-Rank Test indicates that between-cell variability is significantly greater than within-cell variability, for both orientation and spacing ($\Delta\theta$: $p < 0.001$, $\Delta\lambda : p < 0.001$). We perform a number of control analyses to confirm that our results are robust, finding that: (1) the average grid field width does not significantly affect the between-cell variability (Fig. S4), (2) the amount of between-cell variability does not change significantly across the recording duration (Fig. S5), (3) the between-cell variability is not significantly impacted by the boundaries of the arena (Fig. S6), and (4) the amount of between-cell variability is not driven by the presence of cells that have both head direction and grid cell like tuning ("conjunctive" cells [46] – Fig. S7).

In keeping with convention, the reported grid cell properties are the average of those computed for each of the three independent axes in the SAC (Axis 1: aligned to $\approx 0°$; Axis 2: aligned to $\approx 60°$; Axis 3: aligned to $\approx -60°$ [47]). To ensure that this averaging is not contributing to greater between-cell variability, we repeated the analysis above, restricting ourselves to each axis separately. The results again demonstrate that grid properties are significantly more robust within-cell than between-cell (Fig. 4A, B). For each axis, the average between-cell variability for every cell was significantly higher than the average within-cell variability (Fig. 4C, D), as reported by the Wilcoxon-Signed-Rank Test for orientation ($p < 0.001$, for all three axes) and for spacing ($p < 0.001$, for all three axes).

Having demonstrated that grid cell properties are robustly heterogeneous in a single module, we proceed to analyze the remaining recordings in the data set ($N = 420$ cells across 8 modules; one module in the data set had no cells that passed our criteria) [34]. Although there are differences across recordings, we find larger between- than within-cell variability for grid orientation and spacing is present across all recordings (Fig. 5A, B; 80.0% of grid cells have greater between- than within-cell variability for orientation, $\langle\overline{\Delta\theta}_{\text{within}}\rangle = 1.4°$ and $\langle\overline{\Delta\theta}_{\text{between}}\rangle = 1.9°$; 87.9% of grid cells have greater between- than within-variability for spacing, $\langle\overline{\Delta\lambda}_{\text{within}}\rangle = 1.7$ cm and $\langle\overline{\Delta\lambda}_{\text{between}}\rangle = 3.6$ cm). A Wilcoxon-Signed-Rank Test finds that these differences are significant ($\Delta\theta$: $p < 0.001$; $\Delta\lambda : p < 0.001$). We do not find evidence suggesting that the within-cell variability is influenced by grid score (Fig. 5C, D; linear regression for θ: $R^2 = 0.03$, $p = 0.55$ Wald Test; linear regression for λ: $R^2 = 0.07$, $p = 0.14$ Wald Test).

To understand whether the observed grid cell variability may be a heretofore unknown property of MEC computational models, we trained recurrent neural networks (RNN) to perform path integration. These RNN models have previously been shown to develop grid responses [7, 8, 10] and have been argued to develop continuous attractor network structure [12], making them normative models for MEC function. We find that the RNNs develop grid responses with a distribution of grid spacing and orientation (Fig. S8A–D), much like the neural data we analyzed (Fig. 2). Changing the sparsity of the recurrent layer by increasing the strength of weight decay regularization reduces the variability (Fig. S8E–F).

Taken together, our analysis of large-scale MEC recordings demonstrates that grid cells in the same grid module do not have a single grid spacing and orientation, but instead have a restricted range of values around the module mean. This property is a robust feature that cannot be explained by noise from the recording or fitting procedures and emerges in existing normative models of MEC.

Variability in grid properties within individual modules improves the encoding of local space

The variability of grid cell properties within individual grid modules, while statistically significant, is small in magnitude. Can a computational benefit in the encoding of local space be gained from this level of inhomogeneity? How sensitive might such a computational benefit be to the exact amount of variability present?

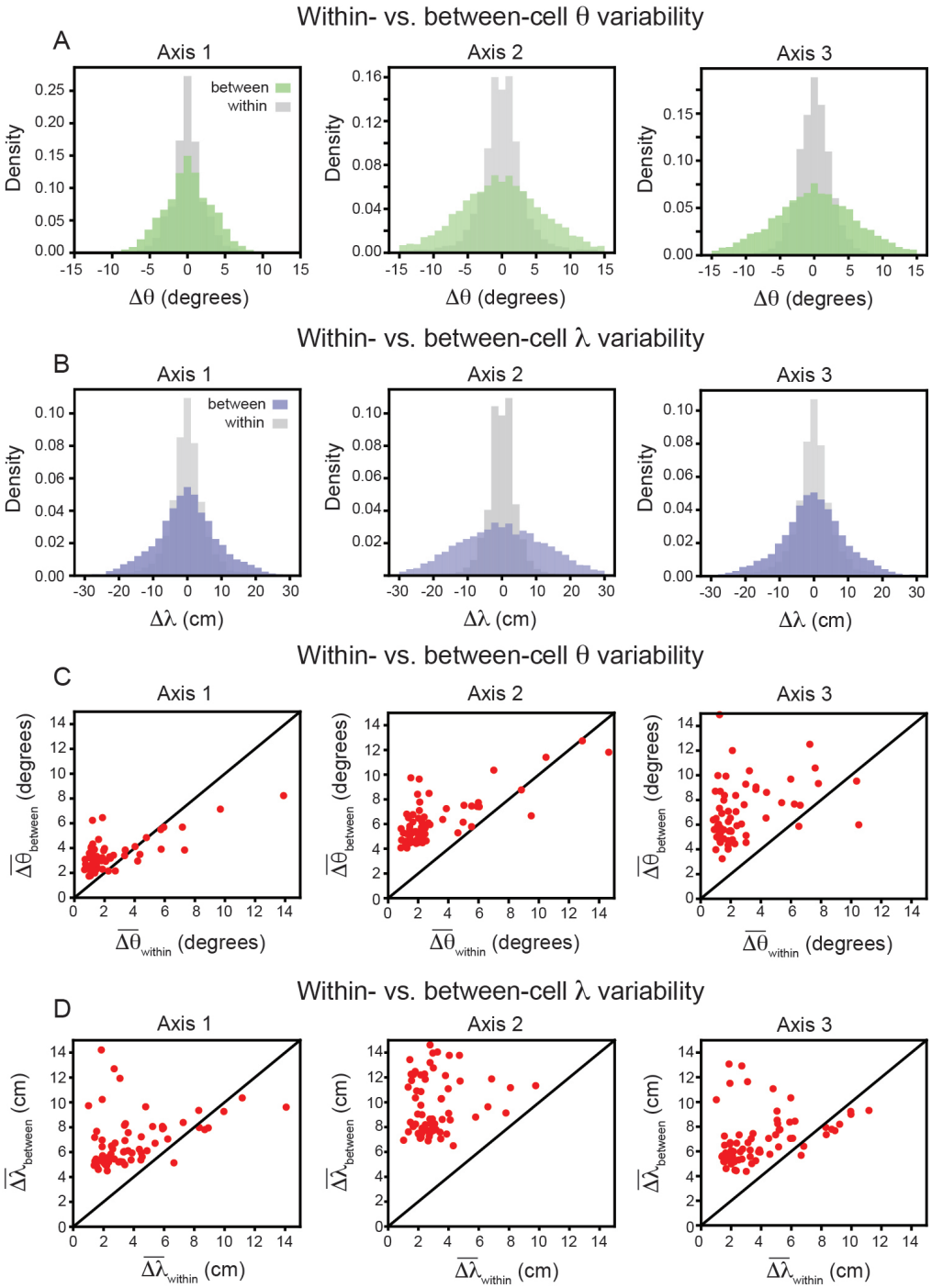

Figure 4: **Variability of grid properties, restricted to the same axis, is a robust feature of individual grid module (recording ID R12).** Same analysis as in Fig. 3 (B) – (E), but for variability measured on each axis independently. For visualization, we exclude a small number of cells that were outside the axes limits, including 2, 5, and 10 cells for Axes 1–3, respectively (C); and 3, 4, and 3 cells for Axes 1–3, respectively (D); these cells were included in non-parametric statistical analyses.

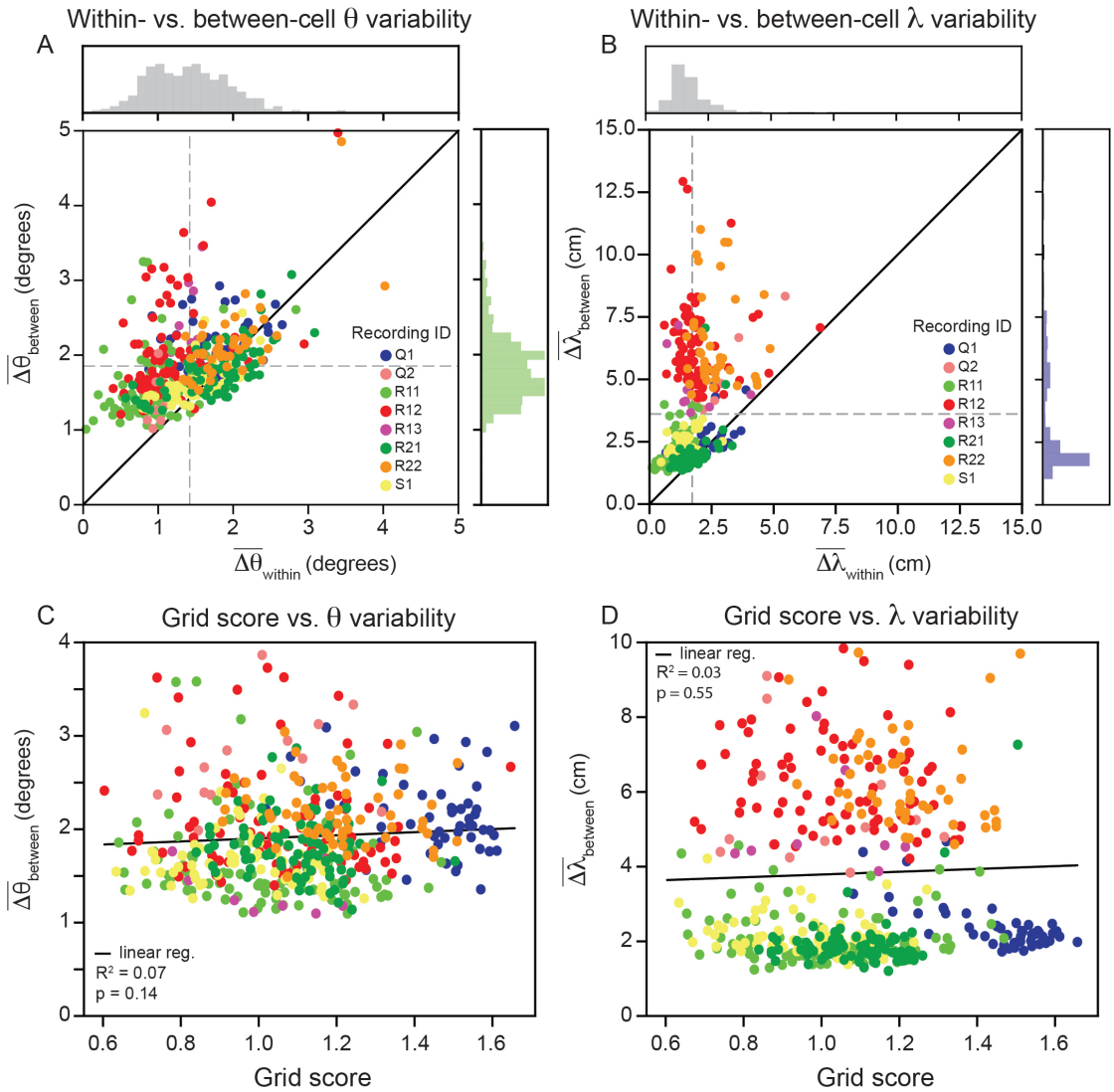

Figure 5: **Within module grid property variability is a robust feature across modules.** (A) Average within-cell variability of grid orientation ($\overline{\Delta\theta}_{\text{within}}$), compared to average between-cell variability of grid orientation ($\overline{\Delta\theta}_{\text{between}}$) for each cell ($N = 420$) across 8 modules (cells colored by their corresponding recording ID). The histogram above the plot shows the distribution of $\overline{\Delta\theta}_{\text{within}}$ and the histogram to the right shows the distribution of $\overline{\Delta\theta}_{\text{between}}$. (B) Same as (A), but for grid spacing. For visualization, 5 cells are excluded ($\overline{\Delta\lambda}_{\text{between}} > 15$ cm), but are included in non-parametric statistical analyses. Dashed gray lines show the population mean. (C)–(D) Average within cell variability of θ and λ (respectively), as a function of grid score. For visualization, 3 and 22 cells are excluded from (C)–(D), respectively, but are included in statistical analyses. Black solid line is linear regression, with R^2 and p-value reported above.

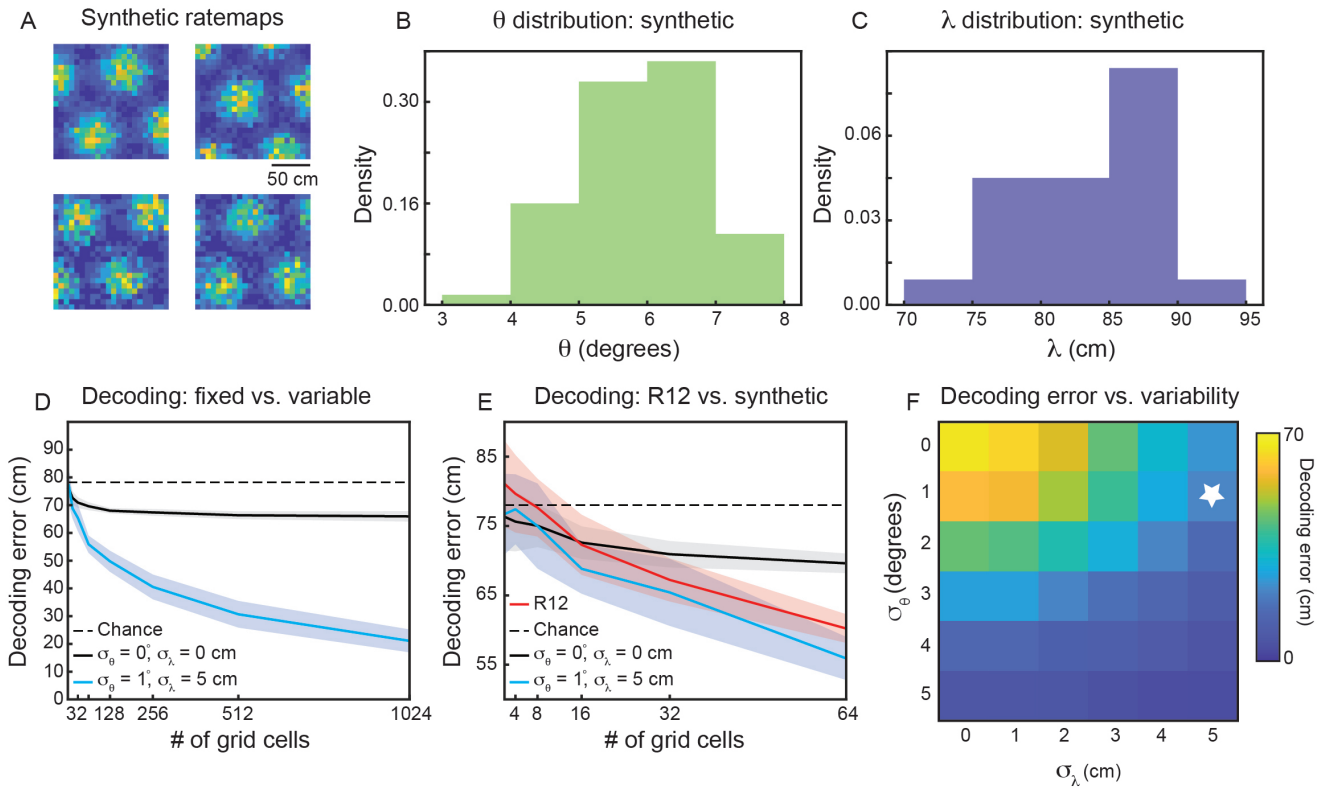

Figure 6: **Variability in grid properties enables improved decoding of local space from the activity of grid cells within a single module.** (A) Example noisy grid cell rate maps generated from a Poisson process. The size of the square arena is set to 1.5 m × 1.5 m to be consistent with what was used in the experimental set-up analyzed [34]. (B)–(C) Distribution of sampled grid spacing and orientation from synthetic population, when using $\hat{\lambda} = 85$ cm, $\hat{\theta} = 6°$, $\sigma_\lambda = 5$ cm, and $\sigma_\theta = 1°$; compare to the distribution measured from real data (Fig. 2E, F). (D) Decoding error, as a function of grid cell population size, with populations having either no variability in grid properties (black line) or variability similar to what was present in the data analyzed (blue line). The solid line is the mean across 25 independent grid cell populations and the shaded area is ± standard deviation of the 25 independent populations. The dashed black line shows chance level decoding error. (E) Decoding error for synthetic populations and real data for up to $N = 64$ cells (red line). (F) Decoding error, over a grid of σ_θ and σ_λ values, for populations of $N = 1024$ grid cells. White star denotes values used in (D).

To address these questions, we generate populations of synthetic grid cells (see Materials and Methods), where we have complete control over the number of grid cells and their firing properties. For simplicity, we assume that all grid cells in our population have grid orientation and spacing sampled from Gaussian distributions, with means $\hat{\lambda}$ and $\hat{\theta}$ and standard deviations σ_λ and σ_θ, respectively. Assigning each grid cell in our population a grid orientation and spacing, we are able to generate "ideal" rate maps [14]. Sampling from a Poisson process on these ideal rate maps, we generate noisy grid cell rate maps (Fig. 6A). Using a simple linear decoder (see Materials and Methods), we can examine how decoding error of local space is affected by the number of grid cells in the population and the amount of variability (σ_λ and σ_θ).

We begin by investigating the decoding capabilities of a synthetic module with properties similar to that of the experimentally recorded module that was analyzed in detail (recording identifier R12; Figs. 2, 3). We therefore set $\hat{\lambda} = 85$ cm and $\hat{\theta} = 6°$. To determine an appropriate value for σ_λ and σ_θ, we subtract the mean between-cell variability by the mean within-cell variability ($\langle\overline{\Delta\theta}_{\text{between}}\rangle - \langle\overline{\Delta\theta}_{\text{within}}\rangle = 2.0° - 1.2° = 0.8°$ and $\langle\overline{\Delta\lambda}_{\text{between}}\rangle - \langle\overline{\Delta\lambda}_{\text{within}}\rangle = 6.9$ cm $- 1.9$ cm $= 5.0$ cm). We interpret these values as the amount of variability in grid spacing and orientation that is not due to noise. Therefore, we set $\sigma_\theta = 1°$ and $\sigma_\lambda = 5$ cm. This amount of variability leads to similar sampled distributions of λ and θ as was found in the real data (compare Fig. 6B, C with Fig. 2E, F).

Applying the linear decoder to populations of synthetic grid cells, we find that decoding error decreases as the number of grid cells is increased (Fig. 6D, blue line). When $N = 1024$, the decoding error is ≈ 20 cm, substantially better than random decoding (Fig. 6D, dashed black line). As expected, this result is not seen in a synthetic population of grid cells with zero variability in grid properties ($\sigma_\theta = 0°$ and $\sigma_\lambda = 0$ cm) (Fig. 6D, black line). In particular, while the restricted number of grid firing fields enables a decoding error smaller than chance, the population with fixed grid properties exhibits little change in decoding error with increasing numbers of grid cells. This demonstrates that the improved encoding is specific to populations with inhomogeneity in their grid spacing and orientation.

To validate that the synthetic population is a reasonable surrogate to the experimentally recorded data, we perform the same decoding analysis on grid cells from recording ID R12 (see Materials and Methods). As the real data is limited in the number of cells, we are only able to compare up to populations of $N = 64$ (Fig. 6E). However, in this restricted range, we find agreement between the synthetic population with grid property variability and the recorded data (Fig. 6E, compare red and blue lines). This supports our hypothesis that the observed amounts of inhomogeneity in grid spacing and orientation can lead to improved decoding of local space.

To determine the extent to which the decrease in decoding error depends the exact variability of grid spacing and orientation, we perform a grid search over 36 pairs of $(\sigma_\lambda, \sigma_\theta)$ values in [0°, 1°, ... , 5°] $\times$ [0 cm, 1 cm, ... , 5 cm]. As expected, when the variability of both grid properties is large, we find nearly 0 cm decoding error of local space (Fig. 6E, bottom right). However, there is additionally a range of σ_λ and σ_θ values that lead to improved decoding, including values smaller than the values matched to the experimental data (Fig. 6E, above and to the left of the white star). We perform the same grid search on synthetic populations of grid cells with different spacings, $\hat{\lambda} = 50$ cm and $\hat{\lambda} = 100$ cm (Fig. S9A, B), finding again that the decoding error drops below that of the fixed population with sufficient grid property variability. For the synthetic module with smaller spacing ($\hat{\lambda} = 50$ cm), the existence of more grid fields in the 1.5 m $\times$ 1.5 m arena leads to greater complexity of interference patterns (Fig. 1), enabling the same amount of variability in grid spacing and orientation to lead to a sharper decrease in decoding error. Additionally, for the synthetic module with smaller spacing, a given value of σ_λ is a larger percentage of $\hat{\lambda}$, as compared to the synthetic module with larger spacing.

Finally, we consider how decoding within a single module compares to decoding across two modules. When the two modules are consecutive (e.g., modules 1 and 2), their mean grid spacing, $\hat{\lambda}_1$ and $\hat{\lambda}_2$, has been experimentally found to be related via $\hat{\lambda}_2 \approx \sqrt{2}\hat{\lambda}_1$ [22]. In such a case, decoding activity from populations with fixed grid properties leads to nearly 0 cm error (Fig. S10A, black line). The addition of variability in grid properties does not significantly change the behavior of the decoding (Fig. S10A, blue line). This suggests that small amounts of inhomogeneity may not disrupt the previously achieved theoretical bounds on decoding from multiple grid modules [24, 25, 28, 48]. However, if the two modules being decoded from are non-consecutive (e.g., modules 1 and 3) then the mean grid spacing can be related by an integer multiple, $\hat{\lambda}_3 \approx 2\hat{\lambda}_1$. In such a setting, the grid fields of the larger module are a subset of the grid fields of the smaller module, up to a rotation (due to the difference in orientation between the two modules), and we again find that variability in grid properties can improve decoding accuracy (Fig. S10B, compare black and blue lines). Similar results are found when using experimentally found values of $\hat{\lambda}$ [22] and not assuming an exact $\hat{\lambda}_{n+1} = \sqrt{2}\hat{\lambda}_n$ relationship.

Taken together, these results suggest that *individual* grid modules can exhibit significant encoding of local space via heterogeneity in their grid properties, even when the extent of the variability in θ and λ is similar to that found in the analysis of the experimental recordings. This benefit can also improve encoding in cases when multiple, non-consecutive modules are considered.

Discussion

The multiple firing fields of grid cells, organized along a triangular lattice, has been historically interpreted as a limiting feature for the encoding of local space. Particularly influential in shaping this view has been

the discovery of the distribution of grid cells into distinct modules, with grid cell spacing (λ) and orientation (θ) preserved within, but not across, modules [22, 23], making integration across multiple modules necessary for spatial information to be decoded [24–28]. While evidence for discontinuity in the grid cell properties across modules is strong, the corollary assumption, that within-module values of λ and θ are identical (up to the bounds of noise), has not been systematically studied.

Analyzing recently collected MEC recordings, we found the range of λ and θ values was large, with examples of grid cell pairs in the same module having over 5° difference in grid orientation and 20 cm difference in grid spacing (Fig. 2). Statistical analysis shows that the variability is more robust than expected from noise, for the majority of grid cells (Fig. 5). This was despite the fact that we used a very conservative criteria for assessing whether a grid cell was consistent enough to be included in the analysis.

Our comparison of within- and between-cell grid property variability was key to our argument, as it was for previous work using it to study the robustness of differences in peak grid field firing rates [41]. The absence of its use in the characterization of distribution of λ and θ [30] may be why the consistency of this heterogeneity was not identified until now. We find that our conclusion holds when performing a number of control experiments (Figs. S1, S4–S7) and whether we treat each grid field independently or take the average across grid fields (Figs. 3, 4). This challenges the assumption that the variability observed in the grid orientation and spacing is attributable solely to measurement noise.

We find that normative recurrent neural network models that develop grid cells when optimized to perform path integration [7, 8, 10] develop similar amounts of grid property variability (Fig. S8). This illustrates the ability of these RNN models to capture aspects of grid cell properties that have not been previously studied, and may provide another explanation for why these models have greater similarity to real MEC recordings than other models [38]. Probing these computational models in greater depth may enable a more detailed understanding of the observed grid property heterogeneity.

We hypothesized that this variability may be used to increase the fidelity at which individual grid modules can encode local space. This idea is consistent with a large body of literature showing that heterogeneity in the responses of populations of neurons increases the robustness of encoding [49–52]. We find, in noisy synthetic populations of grid cells, that a level of variability in grid properties similar to what is quantified in the real data can be sufficient to accurately decode information of local space (Fig. 6). This benefit is increased with larger numbers of grid cells in the population (Fig. 6D) and is observed over a range of values for the underlying variability (Fig. 6F). We find that the improvement was most pronounced in modules with small grid spacing (Fig. S9A), although larger modules can see a decrease in decoding error for amounts of variability consistent with was was found in the analyzed data (Fig. S9B).

We note that our results are additionally aligned with recent findings of heterogeneity in maximal firing rates across individual grid fields [39–41] and grid sheering [36, 37, 42]. Our work further demonstrates that, even in the absence of these perturbations, individual grid modules may encode considerably more local spatial information than previously believed.

Finally, models of the formation of orientation maps in visual cortex have demonstrated that slight angular offsets of retinal mosaics, along which retinal receptive fields are organized, can generate complex patterns [53] similar to those found in visual cortex orientation maps [54]. Our results indicate that grid cells in MEC may take advantage of a similar computational principle, suggesting that mosaic patterns might be a broadly utilized feature of neural coding [55].

Limitations

While the data set we analyzed [34] represents an advance in the ability to simultaneously record from tens to hundreds of putative grid cells, across grid modules, the MEC remains a challenging brain region to access for large-scale neurophysiological experiments. Indeed, with our conservative inclusion criteria, we were ultimately limited by having only 420 grid cells included in our analysis. Future work can perform more detailed and complete characterizations of grid property heterogeneity, as new neurotechnologies that enable larger yield of grid cells are developed [56, 57].

Our decoding analysis, while demonstrating the possibility that variability in grid properties can be used by individual grid modules to enhance the encoding of local spatial information, made several simplifying assumptions, including: (1) independent Poisson noise for neural activity, (2) linear decoding, and (3) normal distribution of grid properties. Although comparison to real data showed that these assumptions are reasonable (Fig. 6E), future work can assess the extent to which these restrictions can be lifted (e.g., to incorporate correlated neural noise), while still enabling individual grid modules to have low decoding error.

Open questions

The heterogeneity in grid properties we characterize motivates the investigation of several new lines of research. Because these are directions that we believe to be fruitful for the field as a whole, we outline them below, with our hypotheses for possible answers.

Q1: How does grid property variability affect continuous attractor network structure? Continuous attractor network models of grid cells [2, 3, 5, 16] enjoy considerable experimental support [23, 30–34], making them one of the "canonical" models in neuroscience. However, these models make use of the assumption that all the grid cells in a given module have the same grid orientation and spacing to simplify the network connectivity. In particular, by assuming equal grid orientation and spacing, it becomes possible to arrange grid cells in a two-dimensional space spanned by their phases (i.e., a "neural sheet" [16]). Neurons close in this space are wired with excitatory connections and neurons far in this space are wired with inhibitory connections. As the data set we analyzed was found by others to provide strong support for the basic predictions of continuous attractor networks (i.e., toroidal topology of the activity manifold) [34], we do not view our results as directly challenging these models. That the RNN models we investigate (Fig. S8) are explicitly trained to perform path integration (and achieve highly accurate performance [12]) supports our hypothesis that variability in grid properties does not necessarily destroy the continuous attractor or path integration capabilities of the MEC. Understanding how this is possible is an exciting future direction, and use of geometric [58] and dynamical systems based tools [59–61] may shed new light on this. We hypothesize that the degree in variability of grid spacing and orientation may strike a balance between being small enough to keep the continuous attractor network structure stable, but large enough to enable encoding of local information of space.

Q2: What causes grid property variability? A natural direction to address is identifying the source of the heterogeneity in grid cell properties we observe. One hypothesis is that this could be driven by "defects" in the specific connectivity pattern that is needed for a continuous attractor. This could be due to synaptic connections between grid cells and other functional classes of neurons in the MEC, such as border [62], band [63], and non-spatial [39] cells. Because the path integrating RNN models have been found to develop analogous responses to these classes [7, 8, 64–66], as well as capture general properties of MEC activity [38], we examined the affect of "connectivity noise" in RNNs. In particular, we swept across different strengths of weight decay regularization, which controls the amount of sparsity in the recurrent layer. We find that smaller weight decay (and thus, denser connectivity between units) led to greater variability in grid properties (Fig. S8E, F). Increasing the weight decay (and thus, enforcing sparser connectivity) led to a reduction of grid property variability (Fig. S8E, F). This supports the hypothesis that the heterogeneity in grid properties we observe may be due to "non-perfect" connectivity.

Alternatively, the coupling between hippocampus and MEC, which has been shown to lead to variability in grid field firing rates (as experimentally observed) [41, 67], may lead to differences in grid orientation and spacing. In particular, a previous computational model that learned grid cells from place cell input, using non-negative principal component analysis, has shown that place field width affects grid cell orientation and spacing [6]. Heterogeneity in the spatial coding properties of place cells has been found along the transverse axis of CA3 [68–70], suggesting there may be a systematic differences in the place field widths of hippocampal inputs to MEC grid cells. Further, it was shown that this place-to-grid cell computational model has a linear relationship between place field width and grid spacing, and a non-monotonic relationship between place field width and grid orientation [6]. These may explain why we find stronger average variability in

grid spacing than grid orientation (Fig. 5A, B).

Q3: How does grid property variability shape hippocampal representations? The projections from MEC to hippocampus suggest that the variability in grid properties may influence hippocampal representations (even if grid cells do not comprise the majority of its inputs [39]). We consider two possible ways in which this may happen. First, given that grid cells have been reported to maintain their grid spacing and orientation across exposures to new environments, while undergoing a change in their grid phase [30, 71], the integration across multiple modules has been necessary to explain place field remapping. However, grid phase plays an important role in generating the specific complex interference patterns that emerge when considering the joint activity of grid cells with variable grid properties (Fig. 1). Thus the reported changes in phase may be sufficient to generate large (and seemingly random) changes in local spatial information conveyed by grid cells to hippocampal cells. This could drive additional changes in the local spatial information projected to hippocampus, as well as explain the significant differences in correlation structure between CA1 neurons across different environments [72].

And second, recent work on hippocampal place field drift [73–77] has demonstrated that there is a significant change in the place field location across time, especially in CA1. One possible source of this phenomenon is the reported instability of dendritic spines on the apical dendrites of CA1 place cells [70, 78–80], ostensibly leading to changes in the MEC inputs to these neurons. However, if grid cells across multiple modules are necessary for local spatial information, turnover in synaptic input is unlikely to cause large changes in the spatial preferences of CA1 neurons, as integration over several scales should provide stable encoding properties. In contrast, different subpopulations of grid cells with variable grid properties can lead to differences in the local spatial information encoded by their joint activity, even if they come from a single module, possibly influencing the spatial preferences of CA1 place cells.

Q4: Why are there multiple modules? Given that our results demonstrate the ability of single grid modules to encode information about local space – a feat previously believed to be possible only if activity from multiple grid modules was integrated together – why is does MEC have multiple modules? While the variability removes the necessity for encoding local space with multiple modules, higher fidelity representations is achievable by integrating across multiple modules (Fig. S10). In addition, the use of grid cells beyond spatial navigation [81–87], where hierarchical representations are important [44, 45], may be a sufficient implicit bias for the formation of multiple modules [88]. In particular, encoding information at multiple distinct scales is critical for multi-scale reasoning, a cognitive function grid cells may support.

Materials and Methods

Electrophysiology recordings

The neural activity analyzed in this paper comes from a publicly available data set, which has previously been described in detail [34]. We provide brief summary of the methodology and the experimental paradigms used during the recordings.

Three male rats (Long Evans – Rats Q, R, and S) were implanted with Neuropixels silicon probes [89, 90]. These probes were targeted at the MEC-parasubiculum region and the surgery was performed as described previously [32, 90]. After three hours of recovery, recordings were performed.

In the recordings analyzed, the rats foraged for randomly dispersed corn puffs in a 1.5×1.5 m^2 square open field arena, with walls of height 50 cm. The rats were familiar with the environment and task, having trained 10–20 times prior to the implantation. The rats were food restricted to motivate their foraging, being kept at a minimum of 90% of their original body weight (300–500 grams).

All procedures in the original study were approved by the Norwegian Food and Safety Authority and done in accordance with the Norwegian Animal Welfare Act and the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes.

Electrophysiology post-processing

The neural activity analyzed in this paper was post-processed, before made publicly available. We describe, in brief, the post-processing performed [34], as well as the post-processing we performed on the downloaded data.

Spike sorting, via KiloSort 2.5 [90], was applied to the data recorded from the Neuropixel probes. Individual units were deemed putative cells if their average spike rate was in the range of 0.5–10Hz, and 99% of their interspike intervals were greater than 2 ms.

For each putative cell, rate maps were constructed by averaging the activity at binned spatial positions in the open field arena. This raw rate map was smoothed, using a Gaussian kernel. The autocorrelation of these rate maps were computed, and a grid score calculated, as described previously [46].

From the downloaded spike trains, we constructed rate maps and autocorrelograms in a similar manner, using code made publicly available [8] (Fig. 2).

Grid module classification

The public data set we analyzed [34] contained the module identity of all putative grid cells. In brief, these module identities were assigned by first projecting the 2D autocorrelogram of every recorded unit onto a 2D space using the non-linear dimensionality reduction algorithm UMAP [91]. Then, the DBSCAN algorithm [92] was used to cluster the units, based on their position in the dimensionally reduced space. Cluster membership served as the basis for grid module classification, with the largest cluster being removed as it was found to not contain spatially selective autocorrelograms. This non-supervised module assignment yielded clusters of high grid scores and similar grid spacing and orientation within each cluster-module. More details can be found in Gardner et al. (2022) [34]. We performed no further analysis regarding module identity.

Computing grid score, orientation and spacing from the spatial autocorrelogram

Grid spacing and grid orientation were computed according to standard methods described in detail previously [22, 93]. Briefly, the goal of the procedure is to identify the location of the six nearest fields in the spatial autocorrelogram (SAC) (Fig. 2A). This was achieved by performing the following steps. First, the SAC was smoothed using a 2D Gaussian filter with $\sigma = 1$. Second, the center of the SAC was excluded by applying a mask. Because the size of the SAC's central peak changes for every module, the radius of such mask was re-computed for each module. Third, we thresholded the SAC by applying an extended-maxima

transform `ndimage.maximum_filter`. Fourth, we identified the center of every field by using the function `scipy.stats.find_peaks`.

Once the peaks of every field had been found, we computed the location of every peak in polar coordinates. We then selected the 6 peaks that were closest to the center of the SAC, based on the computed radial components. Because every SAC is symmetric, we considered for further analysis the 3 peaks closest to the X axis in angular distance (Axis 1, 2, and 3 [47]). Grid spacing was computed as the arithmetic mean of the radial component of the 3 peaks (except when each peak was analyzed separately – Fig. 4). Given that the SAC dimensions are twice of that of the real arena, we multiplied the SAC radial mean by a factor of 2. Grid orientation was computed as the angular mean of orientations (relative to the x-axis) of the three peaks (except when each peak was analyzed separately – Fig. 4).

In order to ensure that subsequent analysis was performed only on cells whose SAC's could be well described by hexagonal structure, we imposed the following constraints: (1) the relative angle between two peaks could not be $< 30°$; (2) the relative angle between two peaks could not be $> 90°$; (3) the values for grid spacing between the peaks could not be substantially different (the ratio of spacings between any two peaks must be > 0.5 and < 2). Cells that did not meet these criteria were determined to have "Poor grid fit" and were rejected from all subsequent analysis. The percentage of all cells that were removed by this inclusion criteria is shown in Fig. S2. We note that cells from one of the nine modules in the publicly available data set that we analyzed had 98.4% of cells rejected by this criteria (recording identifier – R23). We therefore did not include it in any subsequent analysis.

To compute the grid score of recorded MEC cells, we made use of previously published code [8], that is based on metrics that have become standards in quantifying grid cell properties[46].

For analysis of the distribution of grid properties (Fig. 2E, F), we included only grid cells with grid scores greater than 0.85. This was done to demonstrate that variability was present even in cells that exhibit robust grid cell properties. In the subsequent analyses, an alternative criteria is used, which considers the reliability of the grid responses (see below).

Computing spacing from the rate maps

In order to characterize the spacing with a method that is less susceptible to the effects of having grid fields at the boundaries of the environment (as the SAC method could be), we compute the spacing using the rate maps of individual grid cells (Fig. S1A). Once the rate map was computed, we smoothed it with a Gaussian filter, setting $\sigma = 1.5$. Then, using the function `scikit-image.feature.peak_local_max` we extract the position of the center of each firing field. Because our goal with this analysis is to show that the fields in the border do not impact our analysis with the SAC, we restrict ourselves to the 3 peaks that are closest to the center of the arena (Fig. S1A). We compute the distances between those three peaks to each other and report the spacing as the average of the distances measured.

Within and between cells splits

To characterize the within- and between-cell variability of grid spacing and orientation, we employed the following approach. First, we split the data into bins of fixed length (30 seconds). From this, we randomly assigned each interval to one of two blocks (denoted as blocks A and B), with exactly half the total number of intervals in each block. For each grid cell, we computed the grid spacing [$\lambda_A^{(i)}$ and $\lambda_B^{(i)}$] and orientation [$\theta_A^{(i)}$ and $\theta_B^{(i)}$], from the data in each block. The within-cell differences of grid spacing and orientation was determined as $\Delta\theta_{\text{within}}^{(i)} = \theta_A^{(i)} - \theta_B^{(i)}$ and $\Delta\lambda_{\text{within}}^{(i)} = \lambda_A^{(i)} - \lambda_B^{(i)}$. Each grid cell's properties were also compared those of another grid cell, with the match being made using random sampling without replacement. These comparisons were determined as $\Delta\theta_{\text{between}}^{(i,j)} = \theta_A^{(i)} - \theta_B^{(j)}$ and $\Delta\lambda_{\text{between}}^{(i,j)} = \lambda_A^{(i)} - \lambda_B^{(j)}$. This process was repeated 100 times (referring to each iteration as a "shuffle"), per recording. Examples of splits of the data, for different shuffles, is schematically illustrated in Fig. 3A). The resulting distributions of $\Delta\theta_{\text{within}}$, $\Delta\lambda_{\text{within}}$, $\Delta\theta_{\text{between}}$, and $\Delta\lambda_{\text{between}}$, across grid cells and splits of the data were then compared. If the within-cell

variability in grid spacing and orientation was smaller than between-cell variability, we concluded that the variability of grid cell properties was a robust feature of the data and not due to noise.

When performing this shuffle analysis, we found that some cells, despite having a good grid fit when all the data was considered, did not have SACs that were well described by hexagonal structure when the data was split in half. We viewed this a manifestation of unreliable grid coding and a possible confound in our quantification of variability. As such, we introduced a new inclusion criteria (replacing that of requiring a grid score of > 0.85), only considering cells that had poor grid fits on $< 5\%$ of all shuffles. The percentage of all cells that were removed by this inclusion criteria is shown in Fig. S2. In general, we found that cells with high grid score were reliable, although there were exceptions. Additionally, we found that size of the bin used for splitting the data did not significantly affect the percent of cells with good grid fits (passed the prior inclusion criteria) that were considered reliable (Fig. S3).

Synthetic grid cells

To study how variability in grid cell properties might endow the grid code with computational advantages, we generated synthetic grid cell rate maps, so that we could have complete control over the distribution of their properties. These synthetic grid cell rate maps were constructed as follows.

First, the lengths of each dimension the "arena" within which the simulated grid cells exist (L_x and L_y) were set. Then, for $N \in \mathbb{N}$ grid cells, the grid spacing and orientation were sampled via $\lambda^{(i)} \sim \mathcal{N}(\mu_\lambda, \sigma_\lambda)$ and $\theta^{(i)} \sim \mathcal{N}(\mu_\theta, \sigma_\theta)$ where $\mathcal{N}(\mu, \sigma)$ is a normal distribution with mean μ and variance σ^2. Grid phase was sampled as $\phi^{(i)} \sim \mathcal{U}([0, L_x] \times [0, L_y])$, where ϕ is a two dimensional vector, with first component uniformly sampled from $[0, L_x]$ and second component uniformly sampled from $[0, L_y]$. To construct a population with no variability in grid properties (to use as a control), we set $\sigma_\lambda = 0$ m. and $\sigma_\theta = 0°$.

For each grid cell, we generated idealized grid responses by summing three two-dimensional sinusoids [14], such that the activity at $\mathbf{x} = (x, y) \in [-L_x/2, L_x/2] \times [-L_y/2, L_y/2]$ is given by

$$X^{(i)}(\mathbf{x}) = X_i^{\max} \frac{2}{3} \left(\frac{1}{3} \sum_{j=1}^{3} \cos \left[R(\theta_i) \mathbf{k}_j (\mathbf{x} + \phi_i) \pi / \lambda_i \right] \right), \tag{1}$$

where $X_i^{\max}$ is the maximal firing rate, $R(\theta_i)$ is the two-dimensional rotation matrix, with rotation θ_i, and $\mathbf{k}_j$ are the wave vectors with 0°, 60°, and 120° angular differences [i.e., $\mathbf{k}_1 = (1, 0)^T, \mathbf{k}_2 = (1/2, \sqrt{3}/2)^T, \mathbf{k}_3 = (1/2, -\sqrt{3}/2)^T$]. To match the recorded neural data, where individual grid cells have distinct maximal firing rates, $X_i^{\max} \sim \mathcal{N}(13, 8)$. We enforced $X_i^{\max}$ to be within $[2, 30]$, by setting any sampled values outside of this range to the boundary values (i.e., 2 or 30).

To determine the extent to which local spatial information can be decoded from the activity of populations of grid cells with different degrees of variability in their grid properties, we performed the following analysis.

We generated noisy synthetic spike rates of N grid cells by assuming a Poisson process and sampling using the idealized rate maps (Eq. 1). More concretely, the activity of grid cell i at position (x, y) was assumed to be a random variable with a Poisson distribution, whose mean was $X^{(i)}(x, y)$. Thus, the probability of observing $\tilde{X}^{(i)}(x, y)$ spikes, at position (x, y), is given by

$$\tilde{X}^{(i)}(x, y) \sim \mathcal{P}\left(X^{(i)}(x, y) \right). \tag{2}$$

Linear decoding synthetic data

For a given resolution of the arena, we generated $\tilde{X}_j^{(i)}(x, y)$, where $j = 1, ..., 10$. That is, we constructed 10 noisy rate maps. We performed cross-validated decoding by averaging across 9 of the 10 rate maps, to get an average rate map $\hat{X}^{(i)}(x, y)$. For sake of simplicity, consider $\hat{X}^{(i)}(x, y) = 1/9 \sum_{j=1}^{9} \tilde{X}_j^{(i)}(x, y)$. To decode local position from the held out noisy rate map [e.g. $\tilde{X}_{10}^{(i)}(x, y)$], we multiplied the activity at each position by all the positions in the average rate map, taking the sum and assigning the decoded position as that with

the largest value. The Euclidean distance between the decoded position and the true position is considered the error. This was performed 10 times (holding out each rate map once) and the average error across all positions in the environment was then averaged across all 10 of the validation splits.

Linear decoding experimental data

To decode the electrophysiological data [34], we sampled subpopulations of grid cells from the $N = 82$ units (recording ID R12) that passed the criteria described previously. For each subpopulation, we split the recorded activity into 10 evenly spaced temporal bins, and constructed average ratemaps for each bin. This is consistent with what was done in decoding the synthetic data. All ratemaps, for each cell, were normalized by the maximal activity. We then randomly chose 5 of the 10 time bins to serve as the "train" data, the remaining 5 served as the "test" data. The rate maps were averaged and then the linear decoder was applied. We sampled 10 different choices of splitting the data and 25 choices of subpopulation.

Recurrent neural network model

Code used for training and evaluating path integrating recurrent neural network (RNNs) [10, 12] was pulled from `https://github.com/ganguli-lab/grid-pattern-formation`. We used the same hyperparameters as are present on the repository (see Table 1), except when examining the effect of weight regularization on grid property variability, in which case we trained RNNs with weight decay magnitude from $\{10^{-6}, 10^{-5}, 10^{-4}, 2.5 \cdot 10^{-4}\}$. The greater the value, the greater the enforced sparsity. These were chosen as they had previously been found to lead to good path integration performance and have high grid scores [12]. We independently verified that this was true for the resulting RNNs. We computed the grid score, spacing, and orientation using similar code implementations as was used for the neural data. For each value of weight decay, we trained three independent RNNs, using different random seeds.

Table 1: **Parameters used to train RNNs on path integration.** See Sorscher et al. (2019) and (2022) [10, 12] more information on these parameters.

Parameter	Value
Epochs	100
Batch size	200
Batches per epoch	1000
Path length (T)	20
Arena length (L)	2.2 m
Learning rate	10^{-4}
Place cells (n_P)	512
Grid cells (n_G)	4096
σ_1	0.12
σ_2	0.24
Activation	ReLU
Weight decay	10^{-4}
Optimizer	Adam

Acknowledgments

We thank the members of the Goard Lab, Francisco Acosta, Spencer Smith, Caleb Kemere, Will Dorrell, and the DYNS graduate students for useful discussions surrounding this work. We thank Andreas Herz for suggesting the analysis present in Fig. S1 and William Dorrell for suggesting the analysis present in Fig. S9C-D. We thank the eLife reviewers for their suggestions. This work was supported by grants to M.J.G.

from NIH (R01 NS121919), NSF (1934288), and the Whitehall Foundation, and a grant to X.X.W. from NSF (2318065).

References

[1] Torkel Hafting, Marianne Fyhn, Sturla Molden, May-Britt Moser, and Edvard I Moser. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052):801–806, 2005.

[2] Alexis Guanella and Paul FMJ Verschure. A model of grid cells based on a path integration mechanism. In *Artificial Neural Networks–ICANN 2006: 16th International Conference, Athens, Greece, September 10-14, 2006. Proceedings, Part I 16*, pages 740–749. Springer, 2006.

[3] Mark C Fuhs and David S Touretzky. A spin glass model of path integration in rat medial entorhinal cortex. *Journal of Neuroscience*, 26(16):4266–4276, 2006.

[4] Hugh T Blair, Adam C Welday, and Kechen Zhang. Scale-invariant memory representations emerge from moire interference between grid fields that produce theta oscillations: a computational model. *Journal of Neuroscience*, 27(12):3211–3229, 2007.

[5] Jonathan J Couey, Aree Witoelar, Sheng-Jia Zhang, Kang Zheng, Jing Ye, Benjamin Dunn, Rafal Czajkowski, May-Britt Moser, Edvard I Moser, Yasser Roudi, et al. Recurrent inhibitory circuitry as a mechanism for grid formation. *Nature neuroscience*, 16(3):318–324, 2013.

[6] Yedidyah Dordek, Daniel Soudry, Ron Meir, and Dori Derdikman. Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis. *Elife*, 5:e10094, 2016.

[7] Christopher J. Cueva and Xue-Xin Wei. Emergence of grid-like representations by training recurrent neural networks to perform spatial localization. In *International Conference on Learning Representations*, 2018.

[8] Andrea Banino, Caswell Barry, Benigno Uria, Charles Blundell, Timothy Lillicrap, Piotr Mirowski, Alexander Pritzel, Martin J Chadwick, Thomas Degris, Joseph Modayil, et al. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, 2018.

[9] Simon Nikolaus Weber and Henning Sprekeler. Learning place cells, grid cells and invariances with excitatory and inhibitory plasticity. *Elife*, 7:e34560, 2018.

[10] Ben Sorscher, Gabriel Mel, Surya Ganguli, and Samuel Ocko. A unified theory for the origin of grid cells through the lens of pattern formation. *Advances in neural information processing systems*, 32, 2019.

[11] Mikail Khona, Sarthak Chandra, and Ila R Fiete. From smooth cortical gradients to discrete modules: spontaneous and topologically robust emergence of modularity in grid cells. *bioRxiv*, pages 2021–10, 2022.

[12] Ben Sorscher, Gabriel C Mel, Samuel A Ocko, Lisa M Giocomo, and Surya Ganguli. A unified theory for the computational and mechanistic origins of grid cells. *Neuron*, 2022.

[13] Bruce L McNaughton, Francesco P Battaglia, Ole Jensen, Edvard I Moser, and May-Britt Moser. Path integration and the neural basis of the'cognitive map'. *Nature Reviews Neuroscience*, 7(8):663–678, 2006.

[14] Trygve Solstad, Edvard I Moser, and Gaute T Einevoll. From grid cells to place cells: a mathematical model. *Hippocampus*, 16(12):1026–1031, 2006.

[15] Edmund T Rolls, Simon M Stringer, and Thomas Elliot. Entorhinal cortex grid cells can map to hippocampal place cells by competitive learning. *Network: Computation in Neural Systems*, 17(4):447–465, 2006.

[16] Yoram Burak and Ila R Fiete. Accurate path integration in continuous attractor network models of grid cells. *PLoS computational biology*, 5(2):e1000291, 2009.

[17] Licurgo de Almeida, Marco Idiart, and John E Lisman. The input–output transformation of the hippocampal granule cells: from grid cells to place fields. *Journal of Neuroscience*, 29(23):7504–7512, 2009.

[18] John L Kubie and Steven E Fox. Do the spatial frequencies of grid cells mold the firing fields of place cells? *Proceedings of the National Academy of Sciences*, 112(13):3860–3861, 2015.

[19] Daniel Bush, Caswell Barry, Daniel Manson, and Neil Burgess. Using grid cells for navigation. *Neuron*, 87(3):507–520, 2015.

[20] Jake Ormond and Bruce L McNaughton. Place field expansion after focal mec inactivations is consistent with loss of fourier components and path integrator gain reduction. *Proceedings of the National Academy of Sciences*, 112(13):4116–4121, 2015.

[21] Caitlin S Mallory, Kiah Hardcastle, Jason S Bant, and Lisa M Giocomo. Grid scale drives the scale and long-term stability of place maps. *Nature neuroscience*, 21(2):270–282, 2018.

[22] Hanne Stensola, Tor Stensola, Trygve Solstad, Kristian Frøland, May-Britt Moser, and Edvard I Moser. The entorhinal grid map is discretized. *Nature*, 492(7427):72–78, 2012.

[23] Yi Gu, Sam Lewallen, Amina A Kinkhabwala, Cristina Domnisoru, Kijung Yoon, Jeffrey L Gauthier, Ila R Fiete, and David W Tank. A map-like micro-organization of grid cells in the medial entorhinal cortex. *Cell*, 175(3):736–750, 2018.

[24] Xue-Xin Wei, Jason Prentice, and Vijay Balasubramanian. A principle of economy predicts the functional architecture of grid cells. *Elife*, 4:e08362, 2015.

[25] Martin Stemmler, Alexander Mathis, and Andreas VM Herz. Connecting multiple spatial scales to decode the population activity of grid cells. *Science Advances*, 1(11):e1500816, 2015.

[26] Ila R Fiete, Yoram Burak, and Ted Brookings. What grid cells convey about rat location. *Journal of Neuroscience*, 28(27):6858–6871, 2008.

[27] Sameet Sreenivasan and Ila Fiete. Grid cells generate an analog error-correcting code for singularly precise neural computation. *Nature neuroscience*, 14(10):1330–1337, 2011.

[28] Alexander Mathis, Andreas VM Herz, and Martin Stemmler. Optimal population codes for space: grid cells outperform place cells. *Neural computation*, 24(9):2280–2317, 2012.

[29] NM Van Strien, NLM Cappaert, and MP Witter. The anatomy of memory: an interactive overview of the parahippocampal–hippocampal network. *Nature reviews neuroscience*, 10(4):272–282, 2009.

[30] KiJung Yoon, Michael A Buice, Caswell Barry, Robin Hayman, Neil Burgess, and Ila R Fiete. Specific evidence of low-dimensional continuous attractor dynamics in grid cells. *Nature neuroscience*, 16(8):1077–1084, 2013.

[31] Benjamin Dunn, Maria Mørreaunet, and Yasser Roudi. Correlations and functional connections in a population of grid cells. *PLoS computational biology*, 11(2):e1004052, 2015.

[32] Richard J Gardner, Li Lu, Tanja Wernle, May-Britt Moser, and Edvard I Moser. Correlation structure of grid cells is preserved during sleep. *Nature neuroscience*, 22(4):598–608, 2019.

[33] Sean G Trettel, John B Trimper, Ernie Hwaun, Ila R Fiete, and Laura Lee Colgin. Grid cell co-activity patterns during sleep reflect spatial overlap of grid fields during active behaviors. *Nature neuroscience*, 22(4):609–617, 2019.

[34] Richard J Gardner, Erik Hermansen, Marius Pachitariu, Yoram Burak, Nils A Baas, Benjamin A Dunn, May-Britt Moser, and Edvard I Moser. Toroidal topology of population activity in grid cells. *Nature*, 602(7895):123–128, 2022.

[35] Will Dorrell, Peter E. Latham, Timothy E. J. Behrens, and James C. R. Whittington. Actionable neural representations: Grid cells from minimal constraints. In *The Eleventh International Conference on Learning Representations*, 2023.

[36] Dori Derdikman, Jonathan R Whitlock, Albert Tsao, Marianne Fyhn, Torkel Hafting, May-Britt Moser, and Edvard I Moser. Fragmentation of grid cell maps in a multicompartment environment. *Nature neuroscience*, 12(10):1325–1332, 2009.

[37] Julija Krupic, Marius Bauza, Stephen Burton, Caswell Barry, and John O'Keefe. Grid cell symmetry is shaped by environmental geometry. *Nature*, 518(7538):232–235, 2015.

[38] Aran Nayebi, Alexander Attinger, Malcolm Campbell, Kiah Hardcastle, Isabel Low, Caitlin S Mallory, Gabriel Mel, Ben Sorscher, Alex H Williams, Surya Ganguli, et al. Explaining heterogeneity in medial entorhinal cortex with task-driven neural networks. *Advances in Neural Information Processing Systems*, 34:12167–12179, 2021.

[39] Geoffrey W Diehl, Olivia J Hon, Stefan Leutgeb, and Jill K Leutgeb. Grid and nongrid cells in medial entorhinal cortex represent spatial location and environmental features with complementary coding schemes. *Neuron*, 94(1):83–92, 2017.

[40] Revekka Ismakov, Omri Barak, Kate Jeffery, and Dori Derdikman. Grid cells encode local positional information. *Current Biology*, 27(15):2337–2343, 2017.

[41] Benjamin Dunn, Daniel Wennberg, Ziwei Huang, and Yasser Roudi. Grid cells show field-to-field variability and this explains the aperiodic response of inhibitory interneurons. *arXiv preprint arXiv:1701.04893*, 2017.

[42] Gily Ginosar, Johnatan Aljadeff, Liora Las, Dori Derdikman, and Nachum Ulanovsky. Are grid cells used for navigation? on local metrics, subjective spaces, and black holes. *Neuron*, 2023.

[43] Jeff Hawkins, Marcus Lewis, Mirko Klukas, Scott Purdy, and Subutai Ahmad. A framework for intelligence and cortical function based on grid cells in the neocortex. *Frontiers in neural circuits*, 12:121, 2019.

[44] Mirko Klukas, Marcus Lewis, and Ila Fiete. Efficient and flexible representation of higher-dimensional cognitive variables with grid cells. *PLoS computational biology*, 16(4):e1007796, 2020.

[45] Jon W Rueckemann, Marielena Sosa, Lisa M Giocomo, and Elizabeth A Buffalo. The grid code for ordered experience. *Nature Reviews Neuroscience*, 22(10):637–649, 2021.

[46] Francesca Sargolini, Marianne Fyhn, Torkel Hafting, Bruce L McNaughton, Menno P Witter, May-Britt Moser, and Edvard I Moser. Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science*, 312(5774):758–762, 2006.

[47] Tor Stensola, Hanne Stensola, May-Britt Moser, and Edvard I Moser. Shearing-induced asymmetry in entorhinal grid cells. *Nature*, 518(7538):207–212, 2015.

[48] Alexander Mathis, Andreas VM Herz, and Martin B Stemmler. Resolution of nested neuronal representations can be exponential in the number of neurons. *Physical review letters*, 109(1):018103, 2012.

[49] Maoz Shamir and Haim Sompolinsky. Implications of neuronal diversity on population coding. *Neural computation*, 18(8):1951–1986, 2006.

[50] Mircea I Chelaru and Valentin Dragoi. Efficient coding in heterogeneous neuronal populations. *Proceedings of the National Academy of Sciences*, 105(42):16344–16349, 2008.

[51] Julijana Gjorgjieva, Guillaume Drion, and Eve Marder. Computational implications of biophysical diversity and multiple timescales in neurons and synapses for circuit performance. *Current opinion in neurobiology*, 37:44–52, 2016.

[52] Nicolas Perez-Nieves, Vincent CH Leung, Pier Luigi Dragotti, and Dan FM Goodman. Neural heterogeneity promotes robust learning. *Nature communications*, 12(1):5791, 2021.

[53] Se-Bum Paik and Dario L Ringach. Retinal origin of orientation maps in visual cortex. *Nature neuroscience*, 14(7):919–925, 2011.

[54] Gary G Blasdel and Guy Salama. Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature*, 321(6070):579–585, 1986.

[55] Spencer L Smith and Ikuko T Smith. life imitates op art. *Nature Neuroscience*, 14(7):803–804, 2011.

[56] Ryan J Low, Yi Gu, and David W Tank. Cellular resolution optical access to brain regions in fissures: imaging medial prefrontal cortex and grid cells in entorhinal cortex. *Proceedings of the National Academy of Sciences*, 111(52):18739–18744, 2014.

[57] Weijian Zong, Horst A Obenhaus, Emilie R Skytøen, Hanna Eneqvist, Nienke L de Jong, Ruben Vale, Marina R Jorge, May-Britt Moser, and Edvard I Moser. Large-scale two-photon calcium imaging in freely moving mice. *Cell*, 185(7):1240–1256, 2022.

[58] Francisco Acosta, Sophia Sanborn, Khanh Dao Duc, Manu Madhav, and Nina Miolane. Quantifying extrinsic curvature in neural manifolds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 610–619, 2023.

[59] William T Redman, Maria Fonoberova, Ryan Mohr, Ioannis G Kevrekidis, and Igor Mezić. Algorithmic (semi-) conjugacy via koopman operator theory. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 6006–6011. IEEE, 2022.

[60] William T Redman, Juan M Bello-Rivas, Maria Fonoberova, Ryan Mohr, Ioannis G Kevrekidis, and Igor Mezić. On equivalent optimization of machine learning methods. *arXiv preprint arXiv:2302.09160*, 2023.

[61] Mitchell Ostrow, Adam Eisen, Leo Kozachkov, and Ila Fiete. Beyond geometry: Comparing the temporal structure of computation in neural circuits with dynamical similarity analysis. *Advances in Neural Information Processing Systems*, 36, 2024.

[62] Trygve Solstad, Charlotte N Boccara, Emilio Kropff, May-Britt Moser, and Edvard I Moser. Representation of geometric borders in the entorhinal cortex. *Science*, 322(5909):1865–1868, 2008.

[63] Julija Krupic, Neil Burgess, and John O'Keefe. Neural representations of location composed of spatially periodic bands. *Science*, 337(6096):853–857, 2012.

[64] Vemund Schøyen, Markus Borud Pettersen, Konstantin Holzhausen, Marianne Fyhn, Anders Malthe-Sørenssen, and Mikkel Elle Lepperød. Coherently remapping toroidal cells but not grid cells are responsible for path integration in virtual agents. *Iscience*, 26(11), 2023.

[65] Markus Borud Pettersen, Vemund Sigmundson Schøyen, Anders Malthe-Sørenssen, and Mikkel Elle Lepperød. Decoding the cognitive map: Learning place cells and remapping. *bioRxiv*, pages 2024–03, 2024.

[66] Markus Pettersen, Vemund Sigmundson Schøyen, Mattis Dalsætra Østby, Anders Malthe-Sørenssen, and Mikkel Elle Lepperød. Self-supervised grid cells without path integration. *bioRxiv*, pages 2024–05, 2024.

[67] Haggai Agmon and Yoram Burak. A theory of joint attractor dynamics in the hippocampus and the entorhinal cortex accounts for artificial remapping and grid cell field-to-field variability. *Elife*, 9:e56894, 2020.

[68] Heekyung Lee, Cheng Wang, Sachin S Deshmukh, and James J Knierim. Neural population evidence of functional heterogeneity along the ca3 transverse axis: pattern completion versus pattern separation. *Neuron*, 87(5):1093–1105, 2015.

[69] Li Lu, Kei M Igarashi, Menno P Witter, Edvard I Moser, and May-Britt Moser. Topography of place maps along the ca3-to-ca2 axis of the hippocampus. *Neuron*, 87(5):1078–1092, 2015.

[70] William T Redman, Nora S Wolcott, Luca Montelisciani, Gabriel Luna, Tyler D Marks, Kevin K Sit, Che-Hang Yu, Spencer Smith, and Michael J Goard. Long-term transverse imaging of the hippocampus with glass microperiscopes. *Elife*, 11:e75391, 2022.

[71] Marianne Fyhn, Torkel Hafting, Alessandro Treves, May-Britt Moser, and Edvard I Moser. Hippocampal remapping and grid realignment in entorhinal cortex. *Nature*, 446(7132):190–194, 2007.

[72] Eliott Robert Joseph Levy, Simón Carrillo-Segura, Eun Hye Park, William Thomas Redman, José Rafael Hurtado, SueYeon Chung, and André Antonio Fenton. A manifold neural population code for space in hippocampal coactivity dynamics independent of place fields. *Cell Reports*, 42(10), 2023.

[73] Emily A Mankin, Fraser T Sparks, Begum Slayyeh, Robert J Sutherland, Stefan Leutgeb, and Jill K Leutgeb. Neuronal code for extended time in the hippocampus. *Proceedings of the National Academy of Sciences*, 109(47):19462–19467, 2012.

[74] Yaniv Ziv, Laurie D Burns, Eric D Cocker, Elizabeth O Hamel, Kunal K Ghosh, Lacey J Kitch, Abbas El Gamal, and Mark J Schnitzer. Long-term dynamics of ca1 hippocampal place codes. *Nature neuroscience*, 16(3):264–266, 2013.

[75] Thomas Hainmueller and Marlene Bartos. Parallel emergence of stable and dynamic memory engrams in the hippocampus. *Nature*, 558(7709):292–296, 2018.

[76] Walter G Gonzalez, Hanwen Zhang, Anna Harutyunyan, and Carlos Lois. Persistence of neuronal representations through time and damage in the hippocampus. *Science*, 365(6455):821–825, 2019.

[77] Can Dong, Antoine D Madar, and Mark EJ Sheffield. Distinct place cell dynamics in ca1 and ca3 encode experience in new environments. *Nature communications*, 12(1):2977, 2021.

[78] Adi Mizrahi, Justin C Crowley, Eran Shtoyerman, and Lawrence C Katz. High-resolution in vivo imaging of hippocampal dendrites and spines. *Journal of Neuroscience*, 24(13):3147–3151, 2004.

[79] Alessio Attardo, James E Fitzgerald, and Mark J Schnitzer. Impermanence of dendritic spines in live adult ca1 hippocampus. *Nature*, 523(7562):592–596, 2015.

[80] Thomas Pfeiffer, Stefanie Poll, Stephane Bancelin, Julie Angibaud, VVG Krishna Inavalli, Kevin Keppler, Manuel Mittag, Martin Fuhrmann, and U Valentin Nägerl. Chronic 2p-sted imaging reveals high turnover of dendritic spines in the hippocampus in vivo. *Elife*, 7:e34700, 2018.

[81] Nathaniel J Killian, Michael J Jutras, and Elizabeth A Buffalo. A map of visual space in the primate entorhinal cortex. *Nature*, 491(7426):761–764, 2012.

[82] Alexandra O Constantinescu, Jill X O'Reilly, and Timothy EJ Behrens. Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352(6292):1464–1468, 2016.

[83] Dmitriy Aronov, Rhino Nevers, and David W Tank. Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature*, 543(7647):719–722, 2017.

[84] Joshua B Julian, Alexandra T Keinath, Giulia Frazzetta, and Russell A Epstein. Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nature neuroscience*, 21(2):191–194, 2018.

[85] Matthias Nau, Tobias Navarro Schröder, Jacob LS Bellmund, and Christian F Doeller. Hexadirectional coding of visual space in human entorhinal cortex. *Nature neuroscience*, 21(2):188–190, 2018.

[86] Niklas Wilming, Peter König, Seth König, and Elizabeth A Buffalo. Entorhinal cortex receptive fields are modulated by spatial attention, even without movement. *Elife*, 7:e31745, 2018.

[87] Sujaya Neupane, Ila Fiete, and Mehrdad Jazayeri. Mental navigation in the primate entorhinal cortex. *Nature*, pages 1–8, 2024.

[88] Rylan Schaeffer, Mikail Khona, Tzuhsuan Ma, Cristobal Eyzaguirre, Sanmi Koyejo, and Ila Fiete. Self-supervised learning of representations for space generates multi-modular grid cells. *Advances in Neural Information Processing Systems*, 36, 2024.

[89] James J Jun, Nicholas A Steinmetz, Joshua H Siegle, Daniel J Denman, Marius Bauza, Brian Barbarits, Albert K Lee, Costas A Anastassiou, Alexandru Andrei, Çağatay Aydın, et al. Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551(7679):232–236, 2017.

[90] Nicholas A Steinmetz, Cagatay Aydin, Anna Lebedeva, Michael Okun, Marius Pachitariu, Marius Bauza, Maxime Beau, Jai Bhagat, Claudia Böhm, Martijn Broux, et al. Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science*, 372(6539):eabf4588, 2021.

[91] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[92] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996.

[93] William N Butler, Kiah Hardcastle, and Lisa M Giocomo. Remembered reward locations restructure entorhinal spatial maps. *Science*, 363(6434):1447–1452, 2019.

Supplemental material

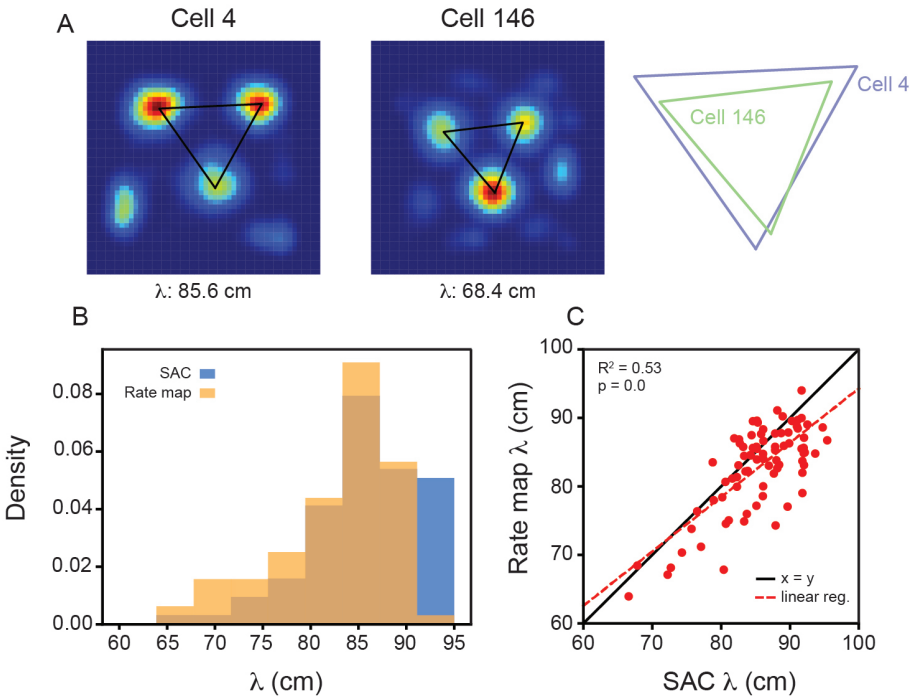

Figure S1: **Variability in grid spacing within a single module exists when computing λ directly from the rate maps.** (A) (Left) Example grid cell rate maps from the same module (recording ID R12) with overlaid triangles, corresponding to the spacing between each of the three most central peaks. Grid spacing is computed as the average of the three lengths. (Right) To aid comparison, the triangles are enlarged (with their relative size fixed) and overlaid. Note that these cells are the same ones plotted in Fig. 2B. (B) Distribution of grid spacing computed using the SAC and the rate maps. (C) The grid spacing of all grid cells, from recording R12, computed using the SAC and the rate map. Red dashed line is the linear regression fit with R^2 and p-value reported above.

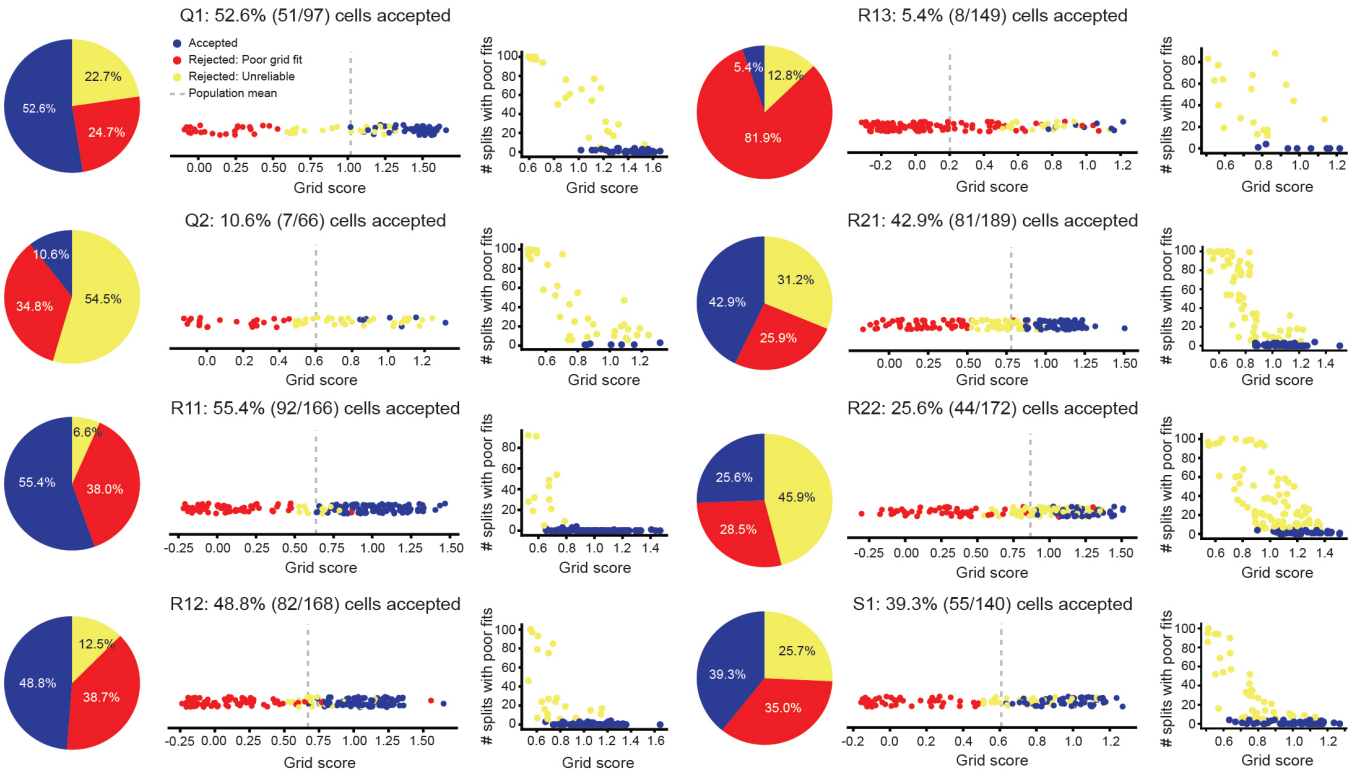

Figure S2: **Accepted and rejected cells across all grid modules.** (Left) The percent of all cells, across all modules, that were rejected by our inclusion criteria. Cells that were rejected as not having SACs, computed with all data, that were well described by hexagonal structure ("Rejected: Poor grid fits") are shown in red. Cells that were rejected as not reliably having SACs, computed from splits of the data, that were well described by hexagonal structure ("Rejected: Unreliable") are shown in yellow. Cells that met these criteria ("Accepted") are shown in blue. (Middle) Grid score of all cells, with coloring denoting whether they were accepted or rejected. Dashed gray line denotes population mean. (Right) The number of splits of the data (out of 100) that cells had SAC's with poor grid fits, as a function of each cell's grid score.

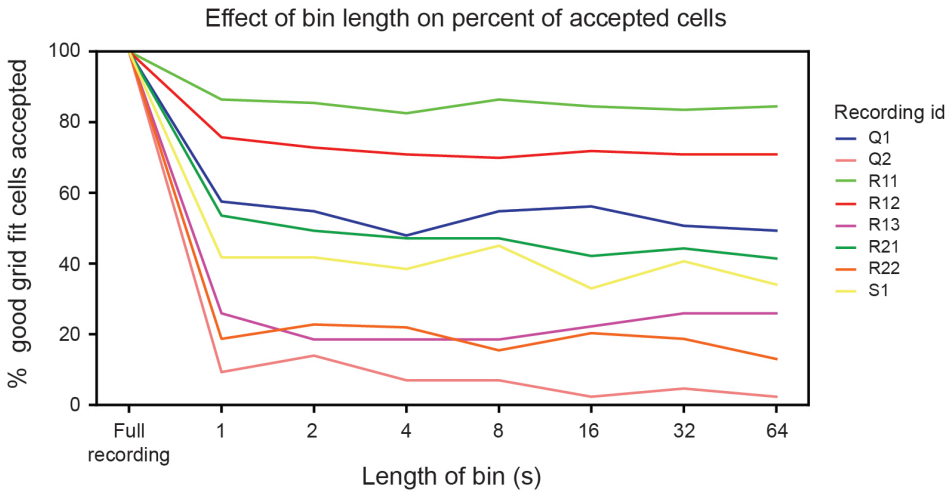

Figure S3: **Bin length does not affect the percent of cells accepted for analysis.** The percent of cells with good grid fits (i.e., those cells that do not get rejected for having "poor grid fits") that are accepted by not being deemed unreliable, as a function of the size of the bins used in the shuffle analysis. All modules are plotted (each line is colored based on its recording ID).

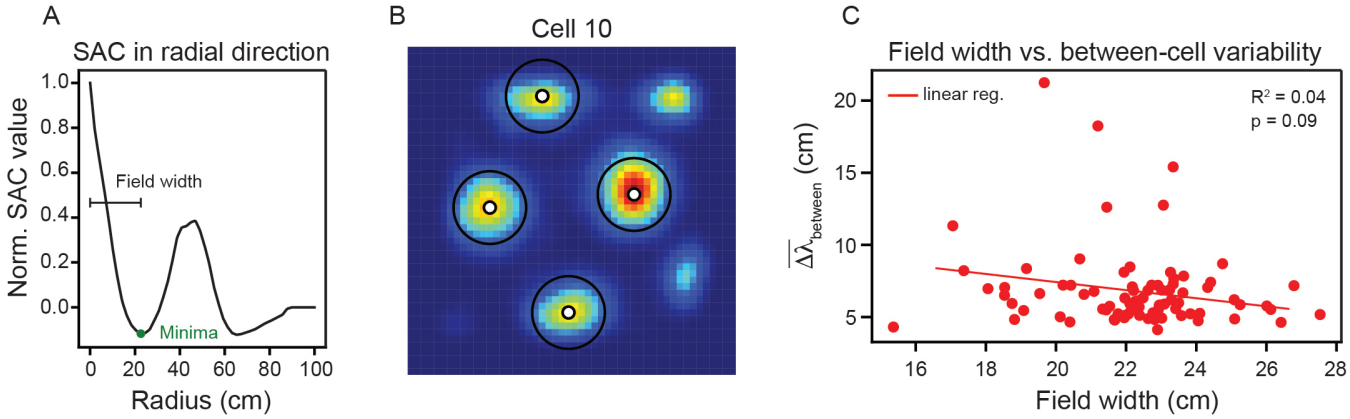

Figure S4: **Average field width does not play a significant role in explaining the between-cell variability in spacing.** (A) Schematic illustration of how the average field width was approximated. The first minima along 1-dimensional slices through the center of the SAC are used to estimate the field width. (B) An example ratemap with the four strongest grid fields overlaid with circles having the diameter of the estimated field width. (C) Between-cell spacing variability ($\overline{\Delta\lambda}_{\text{between}}$) as a function of average field width for all cells from recording R12 that passed our inclusion criteria. Red line is the linear regression fit with R^2 and p-value reported. R^2, for all other modules: Q1 = 0.06; Q2 = 0.40; R11 = 0.04; R13 = 0.03; R21 = 0.12; R22 = 0.12; S1 = 0.00.

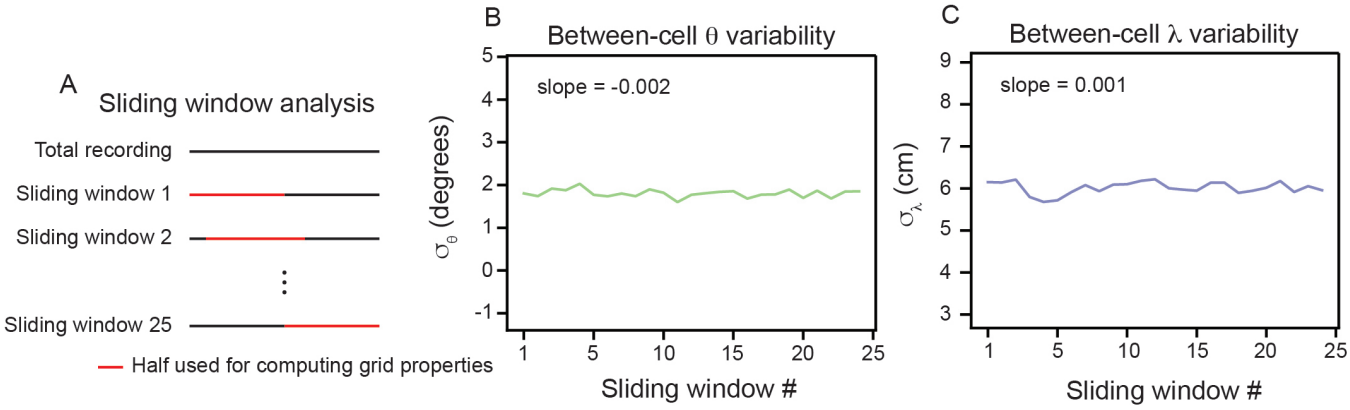

Figure S5: **Between-cell variability does not significantly change across recording time.** (A) Schematic illustrating the sliding window analysis. Red line denotes half of data used to compute the grid properties. (B)-(C) Between-cell variability (estimated by computing the standard deviation of the population θ and λ values) as a function of sliding window number. The slope estimated from a linear regression fit is reported. Slope in θ and λ, for all other modules: Q1 = 0.015 and 0.028; Q2 = −0.039 and −0.003; R11 = −0.001 and −0.005; R13 = 0.005 and −0.000; R21 = 0.014 and 0.020; R22 = 0.044 and 0.016; S1 = 0.034 and 0.014.

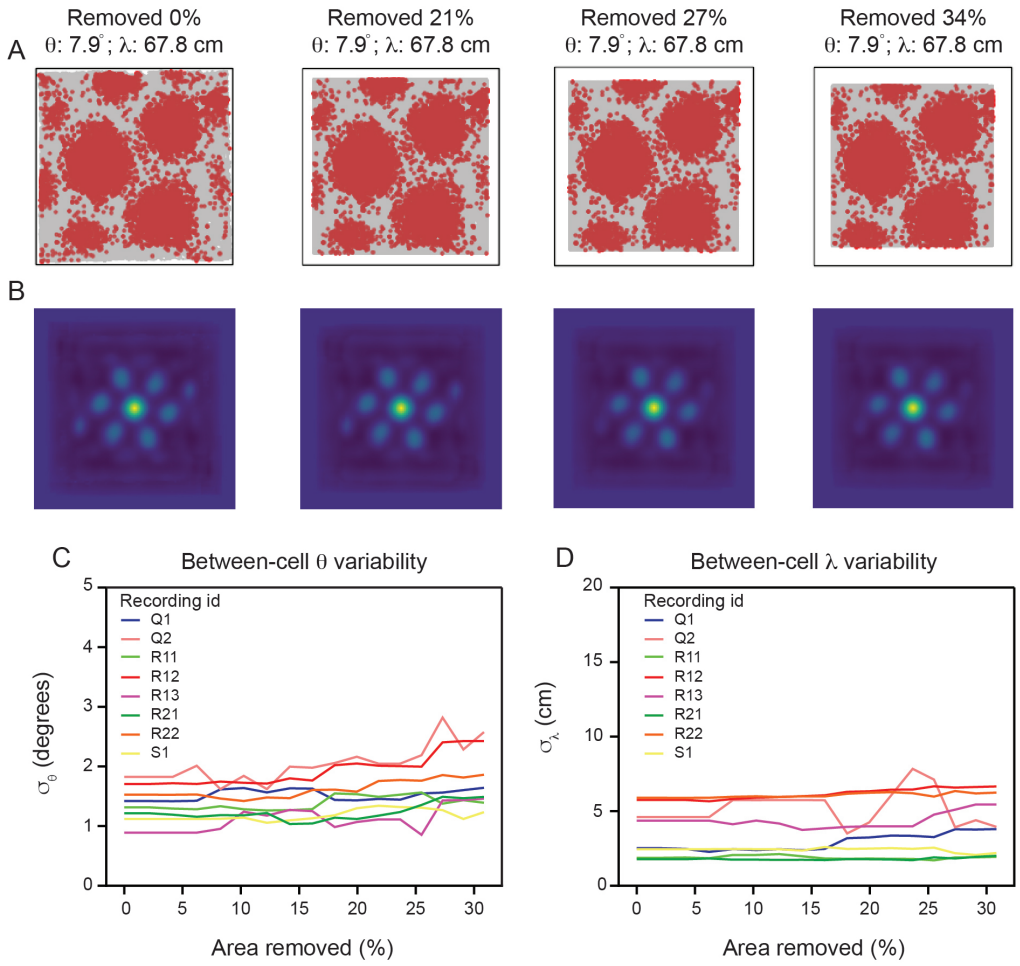

Figure S6: **Variability in grid spacing and orientation is not significantly affected by location of arena boundaries.** (A) Example spike map for one cell (from R12) as increasingly more of the arena is removed. Computed θ and λ are reported above the spike maps. (B) SACs corresponding to the spike maps. (C)-(D) Between-cell variability (estimated by computing the standard deviation of the population θ and λ values) as a function of how much of the arena was removed from our analysis for all modules.

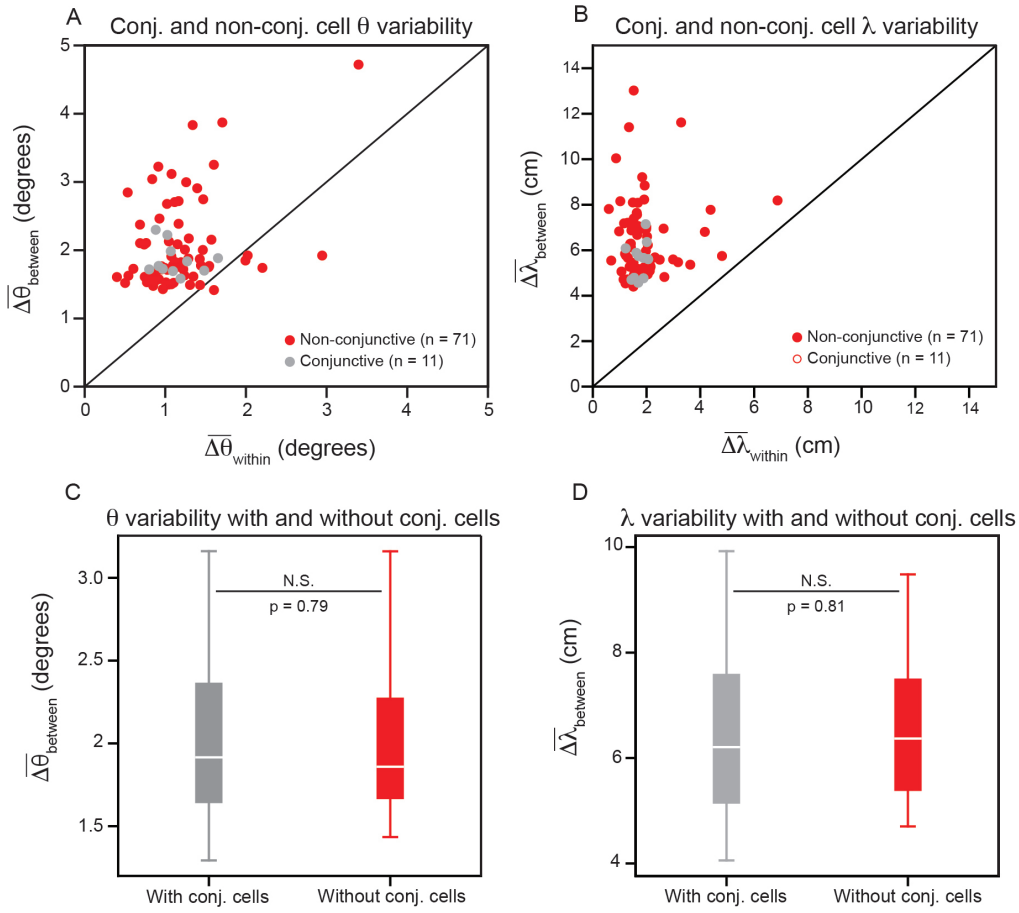

Figure S7: **Conjunctive head direction grid cells do not significantly affect the variability in grid orientation and spacing** (A)–(B) Same as Fig. 2D–E, but with conjunctive and non-conjunctive cells differentiated. Note that the slight difference in position of the dots between this figure and Fig. 3 is due to re-computing with a different random seed. (C)–(D) Distribution of between-cell variability in orientation and spacing, when computing $\overline{\Delta\theta}_{\text{between}}$ and $\overline{\Delta\lambda}_{\text{between}}$ with and without the conjunctive cells. To allow paired comparisons, the variability is calculated for only the non-conjunctive cells, but either includes (grey) or does not include (red) variability comparisons to the conjunctive cells. A paired t-test is used to determine whether the distributions are significantly different (p-values reported in C–D). p-value in θ and λ, for all other modules: Q1 = 0.65 and 0.78; Q2 = 0.86 and 0.92; R11 = 0.13 and 0.75; R13 = 0.99 and 0.97; R21 = 0.29 and 0.98; R22 = 0.87 and 0.95; S1 = 0.07 and 0.77.

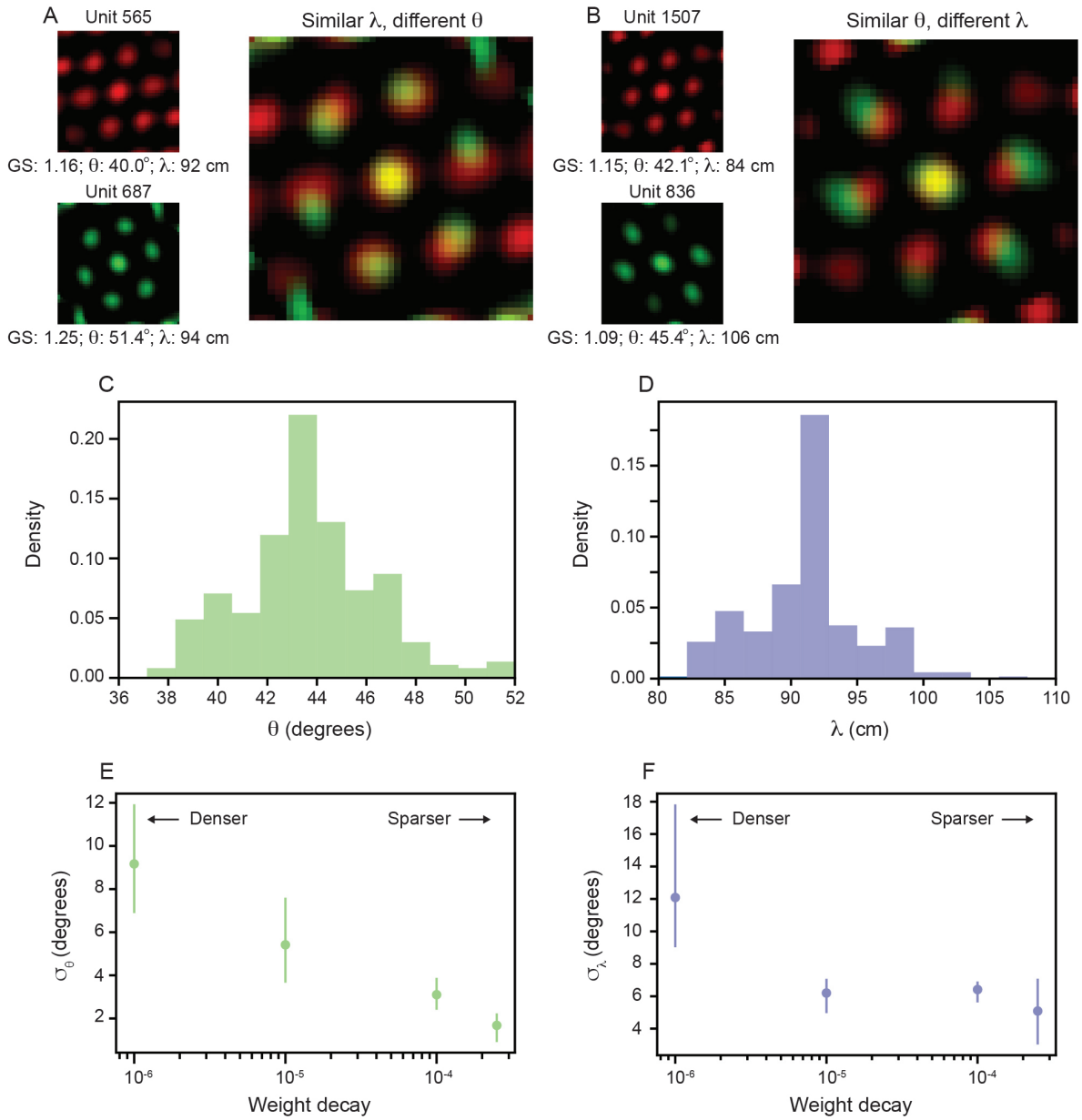

Figure S8: **Path integrating recurrent neural networks (RNNs) that develop grid cells exhibit variability in spacing and orientation that scales with recurrent weight sparsity.** (A)–(B) Spatial autocorrelogram (SAC) overlaid for two pairs of units in the recurrent layer with high grid score; one pair with different θ and similar λ (A) and the other with similar θ and different λ (B). (C)–(D) Distribution of θ (C) and λ (D) across all units in the recurrent layer with grid score > 0.85 (N = 326). Two and four units were excluded for visualization from (C) and (D), respectively, since they were outside the plotting axes. (E)–(F) Standard deviation of θ (E) and λ (F) distributions, σ_θ and σ_λ respectively, across different amounts of weight decay used in training. The circle markers indicate the mean across three independently trained networks and the lines indicate the minimum and maximum of all three networks.

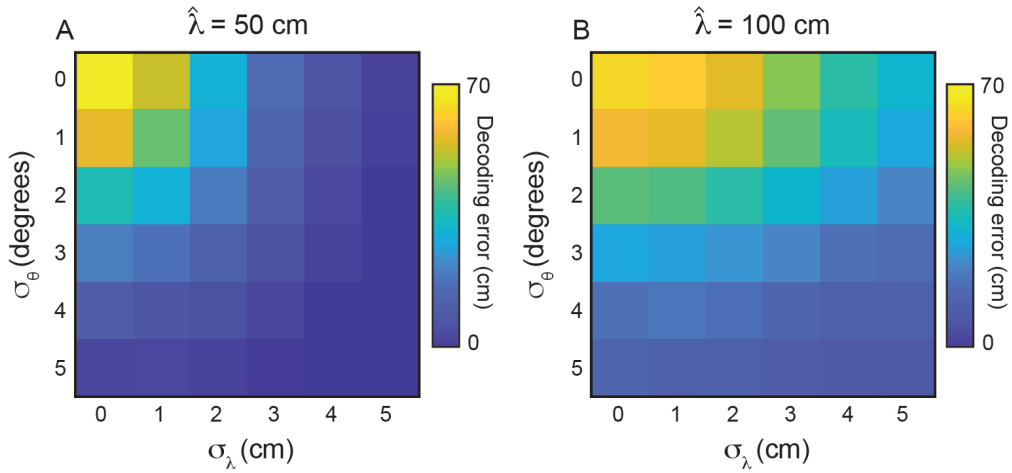

Figure S9: **Variability in grid properties improves decoding of local space for grid modules with different mean grid spacings.** (A)–(B) Same as Fig. 6F, for $\hat{\lambda} = 50$ cm and $\hat{\lambda} = 100$ cm, respectively. $\hat{\theta}$ is set to 0°.

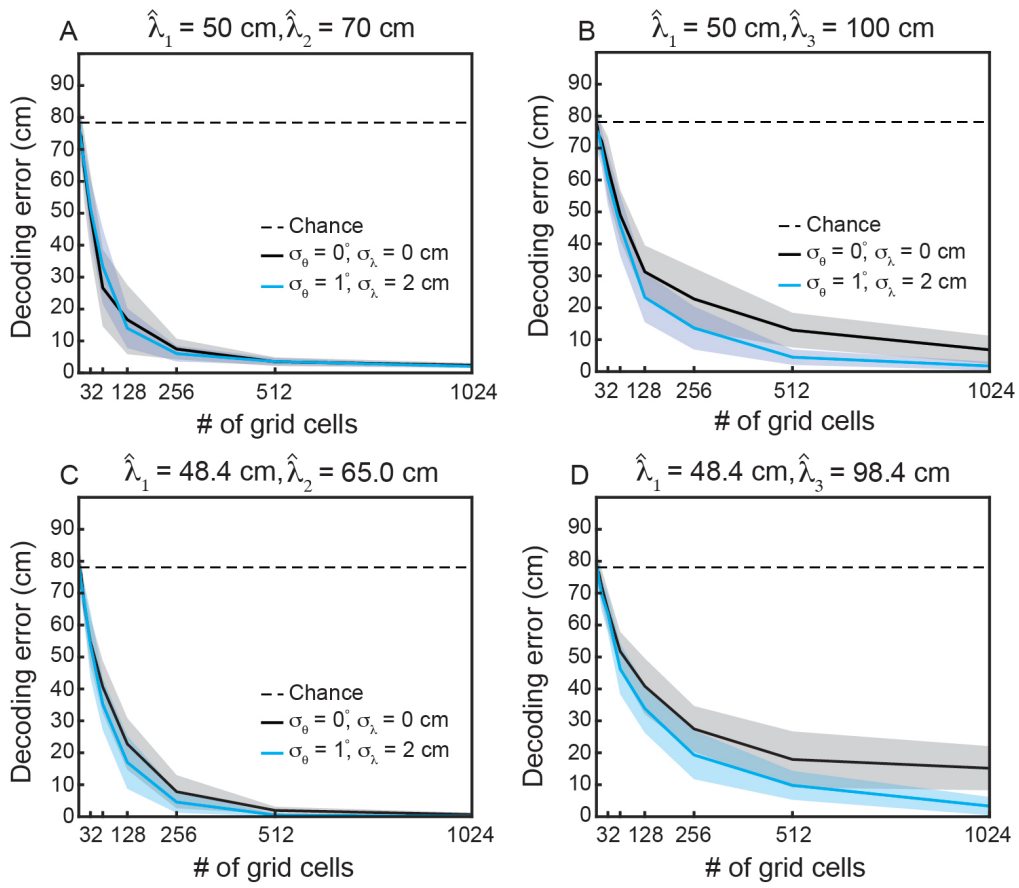

Figure S10: **Variability in grid properties improves decoding of local space for multiple modules, when the modules have integer multiple mean spacing.** (A)–(D) Same as Fig. 6D, when decoding from multiple modules. (A) $\hat{\lambda} = 50, 70$ cm and $\hat{\theta} = 0°, 10°$. (B) $\hat{\lambda} = 50, 100$ cm and $\hat{\theta} = 0°, 10°$. (C) $\hat{\lambda} = 48.4, 65.0$ cm and $\hat{\theta} = 0°, 10°$ (D) $\hat{\lambda} = 48.4, 98.4$ cm and $\hat{\theta} = 0°, 10°$. The values in (A)–(B) were chosen such that $\hat{\lambda}_2 \approx \sqrt{2}\hat{\lambda}_1$ and $\hat{\lambda}_3 = 2\hat{\lambda}_1$. The values in (C)–(D) were chosen to match those found in Rat 14257 from Stensola et al. (2012) [22].